The use of games in citizen science is growing, but can create tension as gaming and science can be seen as incompatible areas of activity. For example, the motivations for winning a game and scientific pursuit of knowledge may be seen as contrary. Over a one-year period, we conducted a virtual ethnographic study of the public forums of two online citizen science projects, Foldit and Galaxy Zoo, the first a project in which gaming is an explicit design feature and the second in which it is not. The aim was to provide a nuanced view of how participants topicalize and respond to tensions between gaming and science. Thematic analysis of discussion forum posts suggests that participants in the two projects respond differently to the tension. By unpacking participant responses to the tension between games and science, our study highlights that citizen science projects using games are not just about fun. To enroll and retain volunteers, these projects also must recognize and manage the implicit normative scientific ideals that participants bring with them to a project. We further conclude that ideals of science embraced by citizen scientists appear to influence the reasons why they participate, either emphasizing equality, like in Galaxy Zoo, or meritocracy, like in Foldit.

Keywords: Citizen science; engagement; games with a purpose; scientific ideals

Introduction

Citizen science is now an accepted term for a range of practices that involve members of the general public, many of whom are not trained as scientists, in collecting, categorizing, transcribing, or analyzing scientific data (Bonney et al. 2014). These practices feature tasks that can be performed by people (e.g., making an observation) or can be performed by people with the support of computational means to organize these efforts (e.g., classifying an image). Games are one of these means. Citizen science games are one form of “Games with a Purpose” (GWAPs) (Law and von Ahn 2011), whose aim is to harness the skills of volunteers for solving scientific problems or for contributing to action projects where citizens intervene in social concerns. Although game-like features such as competition and playfulness have existed in some early projects such as The National Audubon Society’s Christmas Bird Count (www.audubon.org), employing digital games to collect data, classify images, or even solve major scientific problems is a recent but growing phenomenon in citizen science.

Research on the use of games in citizen science is relatively novel but expected to grow rapidly. A central issue is the convergence between gaming and science that can stir controversy as the two can be seen as incompatible areas of activity. For example, the competitive structure of a computer game can potentially be seen to corrupt the quality of data as the motivation to win a game may be incompatible with the scientific pursuit of knowledge. Another risk is that players may focus on fun elements and ignore, neglect, or even cheat on embedded science tasks to get them over with quickly (Prestopnik et al. 2014). In addition, with the advent of digital games in citizen science, tensions between gaming and science can be experienced in a variety of ways by amateur participants.

This study addresses the following research question: How do volunteer participants in online citizen science projects topicalize and respond to the tension between gaming and science? We examine two dissimilar projects, Foldit, where gaming is an explicit design feature, and Galaxy Zoo (GZ), where it is not. The aim is to provide a nuanced view of how amateurs come in contact with the tension. The focus of the article is largely on participants in the gamified project (Foldit) with data from participants in the non-gamified project (GZ) used to foreground contrasts arising from the different project approaches. The ways that participants topicalize and respond to the tension between gaming and science is connected to their beliefs and values, therefore, we will contextualize our findings in relation to the broader literature on participant engagement and motivation in online citizen science projects with special regard to online citizen science games.
Participant Engagement in Online Citizen Science Projects

Online citizen science projects are conducted entirely via the Internet and participants help analyze large sets of data provided to them by scientists (Holliman and Curtis 2014). In these projects, participants can be invited to perform various tasks in terms of format and level of complexity. Distributed computing projects, such as SETI@Home and Folding@Home, in which participants use project software that downloads and analyzes radio telescope data or run software to simulate protein folding, can be seen as amongst the earliest examples of online citizen science. Described by Grey (2009) as “volunteered computing,” these earliest projects require participants only to install software on their computers. Unlike these forms of passive involvement, Grey defined “volunteered thinking” as online projects where participants are engaged at a more active cognitive level. Galaxy Zoo, along with a small number of scientific projects embedded in online multiplayer computer games such as Phylo (Kawrykow et al. 2012), EteRNA (Lee et al. 2014), and Foldit (Cooper 2014), and gamified applications like Flora Caching (Bowser et al. 2014), are examples of volunteered thinking.

Defining the features that may be accepted as parts of games and those that may be rejected is difficult (Elias et al. 2012). However, it is possible to identify features that are present in most, but not necessarily all, games. These include narrative context, feedback, reputations, ranks and levels, competition under rules that are explicit and enforced, and time pressure (Read and Reeves 2009). Salen and Zimmermann (2006) state that a game, in contrast to play, is characterized by rules and artificial competition among players to achieve quantifiable outcomes. While well crafted, this definition does not clearly identify what a game is. Consider science, for example, is it a game under this definition? After all, as Van Noorden (2016) argued, for a hard-pressed post-doc, research can feel like a game where the cards are stacked against them. Similarly, we could follow Elias et al. (2012) and consider “artificial competition” to be something that people do for fun and not for serious purposes such as science or work. However, this approach would entirely exclude the possibility of gaming within science. Since von Ahn (2006) developed the idea of “games with a purpose” (GWAP), games have been used to motivate and sustain crowd participation in and out of science. Both games and elements of games or “gamification,” where game-related ideas are applied to non-game contexts (Deterding et al. 2011), are used in citizen science to develop applications that invite citizens to collect data, annotate images or documents, or solve difficult scientific problems. Several GWAPs developed to support citizen science projects in various disciplinary areas, ranging from biology and biochemistry to astronomy, are described in Lafourcade et al. (2015). Game-based projects such as Phylo and Foldit invite citizens to perform macrotasks involving challenging problems whose solution might take weeks or even months of effort from qualified experts. The goal of involving the general public to accomplish macrotasks is to have a large pool of candidates from which to identify a few talented individuals whose skills can solve challenging problems (Good and Su 2013). On the opposite side, game-based projects such as MalariaSpot (Luengo-Oroz et al. 2012) solicit citizens to perform microtasks that can be accomplished quickly and easily by anyone who can read simple instructions.

Researchers have pointed to games and gamification as providing opportunities to address two common problems in citizen science projects: The lack of resources and the recruitment and retention of participants. They note that features such as badges and competition could support volunteer motivation and retention (Newman et al. 2012; Bowser et al. 2014). Engagement in online citizen science projects including games has only recently begun to be examined systematically, with researchers focusing on the relationship between motivation and engagement. An inverse relationship between desire to contribute and fun was found in a study of a sample of Galaxy Zoo volunteers conducted by Raddick et al. (2010). Volunteers who chose contribution to science as their primary motivation were more likely to contribute greatly, while those who chose fun as their primary motivation were more likely to contribute less. Contributing to scientific research and interest in science were also highly rated by the sample of Foldit players in Curtis’s (2016) study. Intrinsic interest in the scientific project was also found to be the main reason to participate by Iacovides et al. (2013) in their study of eight volunteers (four from Foldit and four from Eyewire). In their recent study of the relationship between crowdsourcing, motivation, and engagement in Eyewire, Tinati et al. (2017) found that despite the use of gamification in the citizen science task, intrinsic motivations including aiding a beneficial cause, advancing scientific knowledge, and learning were the most highly rated reasons to participate. However, the use of gamification within citizen science has also been shown to be task and domain specific (Eveleigh et al. 2013 and Prestopnik and Tang 2015), and may have negative effects on sustaining engagement. When Tinati et al. (2017) used the term “despite” (emphasis in original) to indicate that intrinsic motivations were the most highly rated reasons to participate, in spite of the use of games, they may refer to the normative desirability of using games in science questioned by some critics, who have warned against the use of the crowdsourcing model of research with its potential to cause harm to participants and manipulate the participant into continued participation (Grabber and Graber 2013). Of those who employ games in citizen science and have started to examine their impact on data quality, Prestopnik et al. (2014) found that different reward systems and gamification approaches can influence player recruitment and retention, but do not necessarily affect data quality adversely.

When games are used in citizen science, their elements of motivation must be balanced with the need for relevant scientific outcomes (Cooper 2014). The use of games in citizen science suggests that game designers tend to rely on players’ civic zeal to trigger the excitement of “participatory science” and the rewarding feeling of achieving something in a field considered prestigious by
the general public (Lafourcade et al. 2015). When there are no hobbies to leverage in a project (birdwatching is an example of a hobby leveraged in citizen science), games can produce the same sense of fun and enjoyment as "real world" activities while still successfully linking participants to data collection efforts (Prestopnik and Crowston 2012). While some studies have found the prevalence of intrinsic reasons in motivating volunteers in online citizen science projects, other studies suggest that extrinsic reasons drive engagement as well, especially in a game-based project (Bowser et al. 2013, Rotman et al. 2012). These different results should not be seen as contradicting each other as they reflect the diversity of online citizen science projects and of people participating (or not participating) in them (Tinati et al. 2017). For example, some groups who might not be inspired by the traditional motivations of citizen scientists such as "contributing to the public good" may be willing to engage with a gamified citizen science application. Furthermore, being motivated by intrinsic reasons may not be contradictory to play and fun, as Greenhill et al. (2016) found in their study of the online citizen science platform Zooniverse.org. Their results indicate that "gamified activity" is a form of intrinsic motivation adding a sense of play to work and tasks (Greenhill et al. 2014).

Given the small number of previous studies on participant engagement in online citizen science projects, further work examining whether participants perceive a tension between contributing to science and participating in a game-based project is warranted. In this study, which seeks to contribute to a better understanding of this issue, the broader literature on participant engagement and motivation in online citizen science projects serves to situate the results within the growing citizen science research discourse.

**Brief Background to the Projects**

The two projects chosen for this study are well known in citizen science and have established online communities of participants. Released in 2008, Foldit (http://www.fold.it) is one of the first online citizen science games to be developed. Employing microtasks, it gives players access to a small number of extremely challenging puzzles. Participants are asked to solve protein folding, one of the hardest computational problems in biology. They can play on their own as soloists, or within a team as evolvers. Researchers analyze the top scoring solutions and determine whether there is a native structural configuration that can be applied to relevant proteins in the real world. People do not need to know anything about biology to play the game. Although a little background can help, as suggested by players in the Foldit Wiki, most players are not biologists.

Galaxy Zoo (http://www.galaxyzoo.org/) is an online environment where a large community of volunteers classifies images of galaxies based on an evolving scheme devised by scientists who use the resulting classifications as part of their studies. Since launching in 2007, well over a million GZ volunteers have examined millions of images. Unlike Foldit, GZ is not explicitly a game, although examples of playful, game-like activity that repurposes the central GZ classification tool can be found discussed on the project forums (Greenhill et al. 2016; Anonymous 2015).

**Data and Methods**

We started this study with prior knowledge from both the literature and prolonged ethnographic engagement with the two projects. From January 2015 through January 2016, we conducted a mixed-method study of the public forums used in the two projects with the help of context analysis strategies to filter out data sample for data analyses. We collected data from the two Foldit public forums, a general forum open to any Foldit related topic and a feedback forum specifically intended to invite suggestions for the project team. These forums contained a total of 12,606 posts at the time of data collection. By contrast, GZ has a far larger and more elaborate forum structure that has migrated platforms over the lifespan of the project. With a total of 34 forums across the two generations of the platform, the discussion related to GZ encompassed 674,350 posts at the time of data collection. To assemble a corpus of posts from the Foldit and GZ forums we used a process known as web-scraping to collect the content and metadata for each post and create a series of spreadsheets for each forum. Because the underlying technologies for the Foldit and two generations of GZ forums are different, we used a variety of tools including Martin Balodis’ WebScraper system, import.io, and the Scrapy Python package to define elements of html content, collect, parse, and cleanup the extracted data. Following the guidelines for conducting ethical research as developed by Ess and the Association of Internet Researchers (2002), we extracted content that is officially, publicly, and permanently archived by the two projects. The discussions on them are not generally of a highly sensitive nature. No password is required to access these public forums (login is required only to post messages), and no site policy prohibits the use of their content without previous consent. However, before starting our study, we sought permission from both the Foldit and GZ project teams to collect data from their forums. We also received permission by individual participants in both projects to quote excerpts from some of their posts. When such permission was denied, we paraphrased the content of the posts. The analysis presented in this paper is an aggregate that draws attention to prevailing themes rather than attending to the practices of any individual participant. Therefore, it can be considered to pose a low risk to those individuals.

**Identifying responses**

Because data extraction resulted in an enormous quantity of posts, to identify potentially relevant discussions we used a concordance program (Imao 2016) to search for posts where specific terms could be found in close adjacency to each other (within a range of five words to the left or right). Based on our research question, the following search term pairs were performed: “rule*” AND “game*”; “science*” AND “game*”; “score*” AND “rank*”; “game*” AND “rank*”; “play*” AND “science”; “game*” AND “not*”;
“classif*” AND “competition.” The choice of search terms was driven primarily by our research question, and only secondarily by the frequency of such words in the forum posts. A word frequency analysis using a stop word list to filter out the most common English language function words run to elicit the vocabulary and phrasing used by the Foldit forum participants generated a skewed distribution of content word frequencies. Most frequent terms in the two Foldit forums included “Foldit” (0.99%), “puzzle” (0.55%), “fold” (0.54%), “game” (0.48%) and “score” (0.45%), while the proportion of the terms “science” and “rank*” was 0.15% and 0.07% respectively. Applying the same process to the Galaxy Zoo corpus produced a list with most frequent terms including “galaxy” (1.51%), “zoo” (0.96%), “like” (0.69%), “interesting” (0.66%), and ‘star” (0.51%). Amongst the search terms, the only terms to appear in the GZ corpus with a frequency of more the 0.01% were “classif*” at 0.37%, “science” at 0.12%, and “play” at 0.01%. As can be expected, the most frequent terms on each forum represent the particular activities of each citizen science project with little overlap, although interestingly, the term “science” appears on the forums with roughly the same frequency.

After a rapid reading of the threads containing instances of adjacency of search term pairs, we chose to focus on the threads retrieved by using the search terms “science” AND “game*”; and “play*” AND “science.” This resulted in the identification of 65 Foldit threads and 11 GZ threads. For GZ, despite a far larger quantity of forum posts in the corpus than Foldit, far fewer posts could be found with the chosen adjacency pairs. This can most likely be attributed to the nature of Foldit, which has been explicitly created as a game, versus GZ, which has not. Examining the identified threads, we then conducted close readings using a strategy aimed at identifying a broad range of opinions and experiences of players about the relationship between science and games through posts that describe experiences and stories reported by participants over time. This search for cases continued until informational redundancy or theoretical saturation was reached. In the case of Foldit, this selection process resulted in 15 threads for further analysis (Table 1).

In the case of GZ, five threads of particular relevance were selected based on the adjacency pair analysis (Table 2).

In narrowing down the number of threads, we did not attempt to obtain a representative sample of participants in the two projects, as we did not collect either direct or indirect demographic information about participants.

Data analysis
We developed a coding frame ourselves using a conventional approach, which is considered appropriate when existing theory or research literature on a phenomenon is still limited (Hsieh and Shannon 2005). We used the actual data to derive codes and categories and to create string links between themes and the data they emerge from. The benefit of this approach is to generate findings grounded in the actual data without imposing preconceived categories or theoretical perspectives. We coded the selected Foldit and GZ forum threads using thematic analysis, according to Braun and Clark’s (2006) guidelines with the assistance of a qualitative data analysis software.

Table 1: Foldit threads selected for the analysis.

<table>
<thead>
<tr>
<th>Thread title</th>
<th>Starting date</th>
<th>End date</th>
<th>Total posts</th>
<th># individual participant posts</th>
<th># Foldit team posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>group sharing and personal scores</td>
<td>05/31/2008</td>
<td>06/05/2008</td>
<td>21</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>puzzle balance and the fun factor</td>
<td>06/25/2015</td>
<td>07/02/2015</td>
<td>27</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>how much information are we expected to share</td>
<td>07/18/2010</td>
<td>03/13/2011</td>
<td>35</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>finding the fun in folding</td>
<td>02/20/2009</td>
<td>05/21/2012</td>
<td>36</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>Admins should cancel and re-submit puzzle 174: More Core . . .</td>
<td>07/28/2009</td>
<td>07/29/2009</td>
<td>12</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>player perspective</td>
<td>02/11/2014</td>
<td>02/13/2014</td>
<td>22</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>random evolver</td>
<td>07/16/2012</td>
<td>04/18/2014</td>
<td>41</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>a roll back to the new client</td>
<td>02/04/2014</td>
<td>02/10/2014</td>
<td>38</td>
<td>35</td>
<td>3</td>
</tr>
<tr>
<td>remove allhands scores</td>
<td>02/14/2012</td>
<td>08/13/2012</td>
<td>15</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>new proposal social structure</td>
<td>07/13/2012</td>
<td>07/16/2012</td>
<td>21</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>PDB export and project aims</td>
<td>10/26/2011</td>
<td>12/05/2011</td>
<td>23</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>proposed change to “Overall” Category</td>
<td>04/28/2014</td>
<td>04/30/2014</td>
<td>13</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>proposed Foldit Privacy Policy</td>
<td>05/02/2012</td>
<td>05/19/2012</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>2009–2011 Fold.it at a glance</td>
<td>04/15/2011</td>
<td>04/19/2011</td>
<td>6</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>real science or just another video game</td>
<td>04/15/2011</td>
<td>04/18/2011</td>
<td>6</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

324 285 39
package, NVivo for Mac. Coding involved two individual researchers who coded the threads identified for each project forum independently. During the coding phase, the two coders exchanged and discussed ideas on their respective datasets to achieve a mutual understanding of the coding frame. To identify initial themes, each individual coder first performed an open coding of their dataset to identify key points, positions, and opinions expressed by participants in relation to the use of games and playful, game-like activity.

With respect to coding and sampling the content used in the analysis, we selected the posts that indicated a clearer topicalization of the relationship between gaming and science.

Thematic analysis of the selected themes resulted in codes that were organized into eight organizing themes and two overarching themes (Table 3).

The different codes used in the analysis of Foldit and GZ forum posts, which were grouped into the eight organizing themes, are presented in the Appendix (Tables 5 and 7 respectively). The percentages of the organizing themes in the forum posts in both projects are presented in the Appendix in Tables 6 and 8. The categories represented by these themes are not mutually exclusive, for example, it is possible for one to acknowledge the importance of sharing to ensure fairness while at the same time being concerned with losing competitive positions. Therefore, rather than a polarization of positions, the themes include a range of different positions.

**Results**

**Community and lack of openness**

These two closely related organizing themes include opinions and experiences about two main aspects. Several posts in GZ expressed the belief that knowledge and tools useful to improving the ways that volunteers perform tasks should be shared with all to support a healthy community. For example, one volunteer expressed a common concern that gaming features related to competition such as rankings would hurt the community stating:

> “The whole merit idea worries me from both a management and implication perspective. We are here out of common interest, to enjoy and discuss and learn from what we see. And enjoy the company of a great community in the process. So why complicate matters and try to make it something that it is not?”

By contrast, in the case of Foldit, this position was often motivated by a perceived need to avoid the system becoming the exclusive playground of an ever-decreasing elite group of players with access to knowledge of the best practices and resources available. Posts referred mainly to sharing recipes, small scripts of computer code that automate some protein folding processes, with all players. Many recipes are shared publicly, but certain groups have recipes that can be used only by members, and individuals can have recipes that they keep to themselves. Players can use recipes to carry out their strategies more easily when solving puzzles, increase their score and rank up. Some posts blamed group-shared recipes for hampering the progress of the community and turning cooperative groups into competitive entities. Recipes closely guarded by groups were reported to create unfairness, which, according to one player who had chosen to stop playing altogether, was “perhaps the worst aspect of the game today.” As one player stated eloquently, “… playing a non-profit game helping science for the benefit of ALL people but using recipes available for only some people playing looks ironic to me.”

This concern seems to be connected to the perceived closed nature of the game, as pointed out by a player who lamented the apparent secretiveness of the Foldit project itself, where “wins” are not openly available and a lack of after-game analysis prevents examination of a win and what was done to make it happen. “Neither the winning results or the method of working is available, players don’t get to see and interact with the top solutions or get a synopsis of process.”

Another aspect of concern in Foldit was the possibility of reusing existing puzzle solutions. This possibility elicited discordant views from players. In contrast to the argument that reusing shared solutions at the beginning of a puzzle could help work out new better solutions more efficiently, if someone else’s work is an entirely better direction than one’s own, concerns were expressed about players loading shared solutions and using them without substantial improvement, or incorporating them as if they were their own without properly acknowledging the author(s) of the solution. “… I think that attribution should
be given to the source of the shared protein because not giving such attribution is plagiarism.”

Reliability of scoring/ranking systems and exploring solutions

Posters on the forums of both projects held discordant views on whether scoring or ranking systems were or could be designed to encourage increased quality in the work of volunteers. In relation to GZ where rankings systems had been briefly tried early in the project lifespan, suggestions from volunteers for adding a leaderboard or similar feature were quickly dismissed by both team members and other volunteers. These responses often included the arguments that ranking systems have a negative effect on the scientific validity of volunteer work and that features like leaderboards are antithetical to the nature of a scientific community. This sentiment on the GZ forums, that rankings reduce scientific validity and are antithetical to the spirit of science, stands in stark contrast to the concerns expressed around scoring on Foldit.

Discourse related to scoring on the Foldit forums took place within a clearly articulated gaming frame, where the notion that ranking performance could be antithetical to the spirit of science is not part of the established repertoires. Instead, without questioning the presence of scoring itself, the details of how scoring was implemented was the focus of much discussion. For instance, some posters thought that only top scoring solutions were exciting to work on and relevant from a scientific point of view, while working on what they called “suboptimal, boring, and scientifically irrelevant solutions” could get players to reach positions in the rankings but “no science, no big thrill either.” This comment suggests that the value of low-scoring solutions was not always clear to players, a thought also pointed out by another poster who expressed the desire to hear more from the Foldit scientists about this aspect and whether people should bother to work hard on a solution if it cannot be top scoring. The potential value of low-scoring solutions was recalled by a poster who mentioned one of such solutions leading to Foldit’s greatest success, the crystal structure of the AIDS-causing Mason-Pfizer monkey virus, a scientific problem that the Foldit players solved in 10 days after having been unsolved for 15 years. This player thus underscored that the scoring system was far from perfect, and the global top scores were not the best gauge of value to science. This claim questioned the reliability of the scoring system (which is a complex feature in Foldit) to explore a large number of design solutions potentially leading to innovative and scientifically sound solutions. The “intentionality” of the scoring system can be unclear to the players as this comment expresses in strong terms, “WHO’S DESIGNS ARE BETTER FOR THE SCIENTIFIC RESEARCH? and what if we stumbled upon an incredible breakthrough but didn’t have it recognized for the effort put into it?” Understanding the relationship between high score and good solutions is crucial but also problematic for the Foldit players. Such a relationship is not always direct. For example, a player reported that Foldit has been plagued by false positives (a false positive is when a solution scores well, but actually does not match the native protein well).

Enjoyment/fulfillment

This category emerged from a range of posts on both the GZ and Foldit forums. In the case of GZ, posts that address issues of enjoyment/fulfillment in relation to gaming

Table 3: Frequency of organizing themes.

<table>
<thead>
<tr>
<th>Organizing theme</th>
<th>Frequency (aggregate number of coding references)</th>
<th>Description</th>
<th>Overarching theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competition and sharing</td>
<td>Foldit: 12% (22), GZ: 12% (6)</td>
<td>Express a tension between producing knowledge and competitive concerns</td>
<td>Values</td>
</tr>
<tr>
<td>Community and openness</td>
<td>Foldit: 31% (58), GZ: 18% (9)</td>
<td>Relationship between gaming and the productivity/health of the community</td>
<td>Values</td>
</tr>
<tr>
<td>Reliability of scoring/ranking systems</td>
<td>Foldit: 14% (26), GZ: 4% (2)</td>
<td>Question whether high-score solutions are always the best for science and express the need to differentiate between soloists and evolvers’ solutions</td>
<td>Values</td>
</tr>
<tr>
<td>Exploring solutions</td>
<td>Foldit: 12% (22)</td>
<td>Assert the importance of exploration to provide diverse solutions</td>
<td>Values</td>
</tr>
<tr>
<td>Lack of openness</td>
<td>Foldit: 3% (6)</td>
<td>Question the perceived closedness of the system</td>
<td>Values</td>
</tr>
<tr>
<td>Enjoyment/fulfillment</td>
<td>Foldit: 15% (28), GZ: 16% (8)</td>
<td>Express frustration at loss of enjoyment/fulfillment as a result of gaming mechanics or practices</td>
<td>Values</td>
</tr>
<tr>
<td>Purpose of project</td>
<td>Foldit: 4% (7), GZ: 25% (13)</td>
<td>Express or question what the project is about</td>
<td>Character of the project</td>
</tr>
<tr>
<td>Relationship between gaming and science</td>
<td>Foldit: 9% (16), GZ: 18% (9)</td>
<td>Emphasize one element over the other, or question the choice of one element over the other</td>
<td>Character of the project</td>
</tr>
</tbody>
</table>
were generally of a hypothetical nature, while in the case of Foldit, they tended to be of a concrete nature tied to specific gaming features of the system. GZ volunteers rarely expressed sentiments indicating that they felt that adding gaming elements would increase their enjoyment/fulfilment. Perhaps somewhat counterintuitively, a commonly expressed notion was that GZ was intrinsically a fun activity, and that gaming features would ruin that fun. Typical volunteer statements on the issue followed a pattern where gaming features were depicted as something ugly, and moderators and scientific team members often made statements that adding gaming features would ruin the fun. When presented with suggestions for the addition of contests or gaming features, one moderator noted:

“Thanks for the thought, but we always like people just to come for the fun and do as much or as little as they can contribute. And we have about 150,000 people clicking away at the moment.”

Here again, “fun” is set in opposition to competition and by extension gaming, revealing an established repertoire within the GZ forum community where gaming features were commonly referred to as reducing enjoyment/fulfilment instead of increasing it. This community norm contrasts directly with the norms of the Foldit community, where gaming features are commonly considered to be a key aspect of Foldit as an activity for achieving enjoyment/fulfilment. The relationship between gaming features and enjoyment/fulfilment or “fun” for the Foldit forum community is most clearly articulated in two threads that discussed the roll-out of a new game client in 2013. This release led to frustration with the software and less enjoyment with the gameplay even among some dedicated and committed players who posted about their dissatisfaction in the forum. Playing for science and having fun were not thought contradictory by participants who volunteer to play the game and do not have their career at stake in the game. As a poster noted, if Foldit no longer had enjoyable elements, then they ceased to be “players” in one respect or the other.

**Competition and sharing**

Concerns were expressed by both Foldit and GZ volunteers about the effects of competition. In the case of GZ, a few volunteers posted calls for the introduction of features that would support competition. During some periods, volunteers even engineered their own ranking systems formed from self-reported metrics based on repurposing of participation data displayed to individual volunteers while logged-in to their accounts. During a discussion related to such a volunteer-created ranking system, one participant proposed the creation of an elaborate system tied to the quality of galaxy classifications made by volunteers. This suggestion was quickly met with a number of negative replies, even from volunteers engaged in creating their own rankings who noted that they were only posting numbers of classifications for fun, and that all participant contributions were equally important:

“In this project we’re all equal even if there are participant that go from astronomers to nothing (like me). It’s not a game. I believe that statistic is important but we are using them principally for joking here.

In this way, without the explicit framing of a game, statements from GZ volunteers most often contained negative perceptions of competition-related features such as points and rankings. Similarly, concerns were expressed by Foldit players with regard to issues of competition and sharing. It seems that having to provide high-scoring solutions and rank up as the focus of playing the game constrained the willingness to share ideas and resources. While some posters expressed the belief that sharing is valuable because it levels the playing field and is integral to science, on the other side, they were also afraid of helping “competitors,” losing advantages, and going down in both the individual and group ranks. Players cooperate and share within their respective groups but can have closely guarded recipes and techniques shared only within groups. Therefore, even when sharing was considered valuable, players decided to share their less critical resources. What to share can become a dilemma as this comment indicates eloquently:

“On the one hand, the whole purpose of Foldit being for the good of the scientific community (…) On the other hand - I find myself rather jealously guarded and protective of my very best techniques and discoveries, while sharing mainly only what I might call my “second-tier” insights, techniques and scripts. . . . On further reflection, however, I know that the people I respect and admire the most in this game are not always those at the tops of the scoring heaps . . . rather, they are those few who excel not only in scoring, but in willingness to teach, and to share, and to assist the less accomplished members of their team, and to foster newbies’ growth.”

Thus far we have noted that claims about the necessity of sharing are concerned with creating a two-tiered system where a few players have access to resources unavailable to others. Along the same lines goes the concern for creating barriers to entry making it difficult for new players to learn the tricks of the trade, while “protecting the old guard” and restricting competition in the game. Reluctance of some experienced players to teach newbies more than just the basics of the game was linked to the control exerted by teammates, who would “watch” what their group members share with non-members.

**Purpose of the project and the relationship between gaming and science**

As the findings related to the previous categories show, there was a significant difference in the ways that Foldit and GZ volunteers expressed their sentiments toward gaming in their respective citizen science projects. In both...
cases, however, the core of the established community norms can be understood as the relationship between gaming features and volunteer understandings of the nature of citizen science as an activity. For the GZ forum community, a clearly articulated norm was that gaming is not compatible with scientific values as expressed by members. Statements by GZ volunteers, especially in relation to the relationship of game-like features and the health of the community, often included arguments for aspiring to a scientific ideal expressed as an activity intrinsically based on openness and sharing where all contributions are valued. For example, one volunteer responded to a post about the relative performance of different community members writing:

"I believe that the zoo is a special place. It’s not important if you’re an expert, a scientist, an astronomer, nothing or whatsoever. Or if you classified 10 or 100000 galaxies. We are a team. Everyone is important and no one is predominant. It’s a sort of one for all and all for one. We’re part of a great scientific project not a game."

Participants commonly described each other as being of equal value to the project with scientific contribution and personal learning as their primary goals. The notion that gaming was incompatible with the nature of GZ as a citizen science project was also regularly expressed by moderators and scientific team members with statements during discussions of rankings or other game related features such as:

"this is a voluntary science project, not a game or competition. People contribute and learn as much as they want or can. So it really wouldn't help us meet the objectives of the project."

This sentiment of gaming as antithetical to citizen science and more generally scientific practice was widely expressed by volunteers, moderators, and scientific team members. It highlights an established community norm that was relatively infrequently challenged. By contrast, Foldit forum community members expressed a wide variety of different opinions on the game and framed its purpose accordingly.

While posts analyzed from the Foldit forum generally did not display uncertainty about the link between science and the game, uncertainty was sometimes expressed in the form of a question starting a discussion about the quality of the scientific contribution of the game, such as in the thread on whether Foldit was about “real” science or just another video game. The lack of after-game analysis with the Foldit team to examine solutions and their contribution to science led a player to wonder whether the game had something to do with solving scientific problems. Another player echoed this view, claiming that, without proper context and feedback, what they were doing was simply playing a video game for the highest scores, regardless of whether what they were producing made sense. As responses to these comments indicating frustration of not making a real contribution to science, some other players claimed that their role was not that of “amateur scientists hell bent on finding a big solution to some mysterious problem,” but rather that of assisting the research of a “real” scientist to produce large amounts of results.

Alongside the framing of players as contributing without being on par with scientists, the characterization of Foldit as primarily a game was similarly clear in some posts. For example, one player emphasized the recreational character of the project as a fun game, although many participants found enticing the chance of achieving a real solution, or getting lucky and making a breakthrough.

**Overview of the Results**

It is possible to draw a demarcation line in terms of values between Galaxy Zoo and Foldit, as indicated in Table 4. On the one hand, GZ participants report an imperative of sharing, equality, and openness, and prioritize contributing to science over other goals. In this community, participants are concerned with corrupting science if tasks are turned into competitive games, as this is antithetical to the spirit of openness that ought to be part of science. By contrast, Foldit participants report that competition involves a limited sharing of data and methods, and solutions are often developed in closed rooms to reach a competitive advantage over other players. Foldit participants often express their motivation being based on reaching a high score and that fulfilment lies in the competitive structure of the game.

**Table 4: Overview of the results.**

<table>
<thead>
<tr>
<th>Competition and Sharing</th>
<th>Community</th>
<th>Reliability of Scoring/Ranking</th>
<th>Exploring Solutions</th>
<th>Openness</th>
<th>Enjoyment/Fulfilment</th>
<th>Purpose of the Project</th>
<th>Relationship between Gaming and Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>GZ</td>
<td>Sharing without competition</td>
<td>Equality</td>
<td>Certainty: Antithetical to “spirit of science”</td>
<td>Open discussion</td>
<td>Unconstrained</td>
<td>Belonging, contributing</td>
<td>Science</td>
</tr>
<tr>
<td>Foldit</td>
<td>Competition with limited sharing</td>
<td>Inequality</td>
<td>Uncertainty: ranking in relation to scientific progress</td>
<td>Closed and uncertain</td>
<td>Constrained</td>
<td>Compete</td>
<td>High score</td>
</tr>
</tbody>
</table>
Discussion

As diverse as Galaxy Zoo and Foldit are, so too are the normative ideals of science, the role of both hedonic and extrinsic rewards, and the intrinsic values and beliefs held by participants in relation to the tension between gaming and science. While points and rankings were considered to be key features for enjoyment by Foldit players, they were perceived as creating competition but not fun by GZ participants. This finding is consistent with previous studies which found that the same competitive mechanisms can be rewarding for some and demotivating for others (Eveleigh et al. 2013; Bowser et al. 2013; Preist et al. 2014). For example, previous results found that when leaderboards were used by participants to compare with what others were doing, they were motivating for high-scoring participants, but off-putting for low-scoring participants who felt they could not catch up (Preist et al. 2014). This result implies that the competitive cues provided by leaderboards can influence which social norms are made salient to participants. Along the same lines, some players may find competitive scoring mechanisms more engaging than collaborative scoring mechanisms (Siu et al. 2014). The use of a competitive scoring mechanism in Foldit seems to influence the decisions of some players to take risks and try innovative strategies to achieve groundbreaking solutions instead of “playing safe” and achieving unoriginal but possibly high-score solutions.

The importance of sustaining a healthy community in GZ is consistent with the desire to contribute as a primary reason for participation found by Raddick et al. (2010). Our result is also consonant with the orientation toward hedonic, rather than utilitarian, “playful experiences” as forms of social interaction useful to building and maintaining a sense of inclusion (Greenhill et al. 2016). This finding suggests that the instrumentalisation of game-like behavior is where the tension lies in GZ, rather than a rejection of “playfulness” altogether. The findings of this study also point to the value of sharing and openness for sustaining a healthy community, a notion consistent with the importance of a sense of inclusion for both intrinsically and extrinsically motivating and encouraging participants to continue contributing to GZ (Greenhill et al. 2016). In relation to Foldit, our findings resonate with the study of three gamified citizen science projects conducted by Tinati et al. (2016), who found that extrinsic motivations such as having fun and the ability to compete with others, along with intrinsic motivations such as helping to improve scientific knowledge, were important.

Our study adds to the existing literature on participant engagement by suggesting that the implicit normative ideals of science that amateurs have should be taken into account when considering what motivates them. However, our findings also raise questions about the ways that citizen scientists’ accounts of science relate to canonized ideas about values in science. In this regard, the clearest case is what we might call the “Mertonian dilemma” (Merton 1942) of how to regulate fierce competition in science. In the case of Foldit, the volunteers reported a dissonance between open science for everyone on the one hand, and the secrecy needed for achieving a high score in the game on the other. The “jealous guarding” of secret methods and scripts would be a violation of the Mertonian norm of communalism and organized skepticism. As a result, we would have to revise parts of what Franzoni and Sauermann (2014) claim to be a core organizing principle of citizen science, namely its radical openness. Among the GZ volunteers, however, there is a much stronger ideal of science as teamwork, where each individual contributor is equally important in helping science. In this respect, citizen scientists resemble the puzzle-solving researcher portrayed by Thomas Kuhn in his Structure of scientific revolutions (1962). This can be associated with the strong opposition to contests found among the GZ volunteers for whom contributing is never about winning and even the smallest task is part of the larger paradigmatic puzzle of “normal science.” The difference between the values of the Foldit and GZ communities can be thought of as an opposition between equality and meritocracy. This distinction may bring to mind the difference between socializers and achievers in Bartle’s (1996) taxonomy of player types, but we argue that the perspectives underlying these opposing values resonate more with Yee’s (2006) Social component and Achievement component, respectively. Equality and meritocracy are found in both Foldit and GZ, although in different ways. Equality in GZ means that every contributor is equally important for science as a whole because what counts in the end is the collective effort. In Foldit, however, contributors are not equal in their performances. Instead, they are equal with regards to the rules of the game. Everybody, irrespective of social background, can score high and make a qualitatively better contribution in the game compared to his or her opponents. Whether this translates into scientific discoveries is unclear to the volunteers in Foldit. In GZ each contribution is valued as an equal contribution, even if a participant only classifies a handful of galaxies.

A notable result of this study is the problematization of the dualism between games and science. While a sharp separation between these two entities is often implied by participants in both Foldit and GZ, emerging from the emphasis they place on one element over the other, the perception of the “hybrid nature” of Foldit has started to emerge. This result is consonant with Horstman and Chen (2012), who also analyzed Foldit forum data and found that much of what players said around playing Foldit “is representative of a hybrid space between science and gaming practices where a diverse set of scientific identities and understandings of science coexist among the players” (p. 3). This hybridity suggests that GWAPs, especially complex ones like Foldit, are better seen as a cognitive technology participating in the distributed thinking that is an integral part of the problem-solving system we identify with human intelligence and scientific discovery, rather than merely an instrumental component.
Implications of the results
Based on our results we draw the following higher-order implications, which might be considered when designing online citizen science projects:

- Game design decisions (e.g., how a game score is calculated), such as competitive scoring mechanisms, can influence the performance of players in terms of selection of strategies.
- Volunteer contributors develop different sets of values with regard to science, and citizen science project managers need to be sensitive to such value changes within the community of volunteers.
- While games often give instant attribution to the highest score, it is also important to make attribution and credit for volunteer contributions in non-gaming contexts clear and fair.
- Competition may both attract and dishearten volunteer participants to contribute.

Limitations and further work
We consider the major limitations of this study to be in the context of the generalizability of our findings as they are mainly informed by the discussion forum posts from the participants in the two projects. As a consequence, our findings are limited by the scope of the discussion forums and the characteristics and number of citizen scientists who participate in them, particularly a low number in the case of Foldit. In this regard, our results may not provide a fair representation of the total populations of GZ and Foldit participants. However, while several published studies on citizen science focus on the analysis of single systems (Tinati et al. 2017), this work has looked across two platforms to compare and contrast the tensions between gaming and science in two projects.

The practical circumstances under which we conducted this study allowed for single coding of each project forum sample. While we realize that interjudging reliability is often perceived as the standard measure of research quality to validate a coding scheme, some scholars argue that a single researcher conducting all the coding can be both sufficient and preferred in studies where being embedded in ongoing relationships with research participants is critical for the quality of the data collected (e.g., Morse and Richards 2002). In this study, the two single coders are particularly familiar with their respective project, as they have spent time getting to know the technical environment and establishing relationships with participants. Finally, to ensure relevance of our interpretations of the data, we discussed our analysis and the results in several research group settings, as well as the 2017 Citizen Science Association conference.

With respect to the use of a conventional approach to the development of the coding scheme, while it allows generating knowledge based on participants’ unique perspectives and is grounded in the actual data, it also fails to develop a comprehensive understanding of the context, thus failing to identify key categories (Hsieh and Shannon 2005).

There are a number of gaps in our knowledge about the ways in which volunteers perceive gamified citizen science, which are highlighted by our findings and would benefit from further research. Following from this study, we need to better understand the role of normative ideals of science in the cultures of contribution developed by citizen scientists. How do citizen scientists develop scientific values? Arguably, scientists play an important role in this development, as they are often those who design online citizen science projects, define their purpose, and decide how citizen contributions are valued and acknowledged. The competitiveness of a gaming approach seems to be a contested issue among citizen scientists, therefore game-based projects should be carefully considered not only for the opportunities they offer to attract and retain contributors, but also to support responsible research and to strengthen collaboration between science and society.

Additional File
The additional file for this article can be found as follows:

- Appendix. Forum post codes and organizing themes. DOI: https://doi.org/10.5334/cstp.101.s1

Notes
1 The Scopus database retrieved 78 articles when searching for [‘citizen science’ AND ‘gam*’] and the Web of Science returned 57 articles with a similar search string [TS =“citizen science” AND TS=“gam*”]. The oldest articles date from 2010 and the oldest conference proceedings are from 2009. Search conducted in January 31st, 2017.

Funding Information
This research was supported through funding from Marianne and Marcus Wallenberg grant no. 2013.0020.

Competing Interests
The authors have no competing interests to declare.

References


