



DEPARTMENT OF EDUCATION,  
COMMUNICATION AND LEARNING

# SUPPORT, CONNECT, AND ORGANIZE

Exploring moderator roles in citizen science  
discussion forums

**Frauke Rohden**

---

Thesis:	30 higher education credits
Program and/or course:	International Master's Programme in IT & Learning
Level:	Second Cycle
Semester/year:	Spring term 2016
Supervisor:	Marisa Ponti
Examiner:	Johan Lundin
Report no:	VT16-2920-001-PDA699

# Abstract

Thesis: 30 higher education credits  
Program and/or course: International Master's Programme in IT & Learning  
Level: Second Cycle  
Semester/year: Spring term 2016  
Supervisor: Marisa Ponti  
Examiner: Johan Lundin  
Report No: VT16-2920-001-PDA699  
Keywords: moderator roles, online forums, citizen science, virtual organizations

---

**Purpose:** The purpose of this study is to explore how moderators act in the online discussion forums of two big citizen science projects, Galaxy Zoo and Foldit, and to find out which roles and influences moderators might have in these projects.

**Theory:** A „conceptual model of virtual organizations for citizen science” based on ideas from work teams and crowdsourcing projects (Wiggins & Crowston, 2010) is used to contextualize the findings of this study.

**Method:** All posts from the two chosen forums were collected and a subset of moderator posts selected for content analysis. Moderator actions and roles were categorized using a framework from an educational context (Asterhan, 2011) and observations on moderator roles in technical help forums (Frith, 2014).

**Results:** It was found that the moderators on Galaxy Zoo and Foldit have a range of important roles in their respective communities: They connect different stakeholders, organize information from different stakeholders and support participants in many aspects, offering pedagogical support as well as recognition of volunteers' contributions. Several moderator actions might contribute to offering learning experiences and resources for learning to the forum users. Furthermore, citizen science projects seem to include aspects of educational as well as other settings both in their overall setup and in their discussion forums.

# Foreword

Since October 2015, I have had the chance to join the regular meetings of the project group on citizen science research at Gothenburg University. This sparked my interest in citizen science and helped me to choose and narrow down the topic for this thesis. I am very thankful for all the input and support I have received from the group during my research. I especially would like to thank Thomas Hillman for sharing his data from Galaxy Zoo Talk with me and helping me to narrow down topics and large amounts of data, and Marisa Ponti for sharing her insights on Foldit and for her encouragement and support as my thesis supervisor.

Gothenburg, Mai 2016

Frauke Rohden

# Table of contents

Part 1: INTRODUCTION TO THE ARTICLE .....	1
Overview and research idea .....	2
Background .....	4
Citizen science.....	4
Can citizen science projects be conceptualized as communities? .....	5
Literature Review on Forum Moderation.....	7
Studying moderation in online discussion forums .....	10
Roles of users in discussion forums .....	11
Styles of moderation.....	12
Purpose of Moderation .....	12
Moderator Behavior .....	13
Automated Moderation.....	13
Research Design.....	16
Ethical concerns of the data collection.....	16
Findings.....	18
Long posts with several actions.....	18
Moderator roles .....	19
Moderators as information architects .....	19
Moderators as tone-setters.....	20
Moderators as translators.....	20
Moderators as knowledgeable non-technical experts.....	21
Part 2: ARTICLE: Exploring moderator roles in citizen science discussion forums .....	22
Background .....	23
Discussion forums and moderation in citizen science.....	23
Theoretical Framework .....	23
Methods.....	26
Overview of the two citizen science projects .....	26
Galaxy Zoo .....	26
Foldit .....	26
Data collection and sampling .....	27
Data analysis .....	28

Findings.....	30
Context of the two forums.....	30
Moderator actions.....	30
Pedagogical support .....	32
Interaction support.....	32
Managerial support.....	33
Moving forward.....	33
Involved participation.....	33
Social support.....	33
Technical support .....	34
Moderator roles .....	34
Discussion .....	36
Comparing the two forums.....	36
Moderators within virtual organizations of citizen science.....	36
Moderators as individuals .....	36
Connecting the individual and the organizational level .....	37
Finding, collecting and presenting knowledge.....	37
Increasing satisfaction and learning .....	38
Conclusion.....	40
Part 3: DISCUSSION AND CONCLUSION OF THE ARTICLE.....	41
Discussion .....	42
How does citizen science compare to educational and other settings? .....	42
Can discussion forums support learning in citizen science projects?.....	43
How do moderator roles in citizen science projects relate to education?.....	44
Moderators as learners.....	44
Moderators as educators .....	44
Challenges, limitations, and further research opportunities .....	45
Summary and Conclusion .....	47
References .....	48
Appendix 1 .....	52

## **Part 1: INTRODUCTION TO THE ARTICLE**

This thesis aims to explore aspects of moderation in the online forums of two citizen science projects, Galaxy Zoo and Foldit. It is examined how the forum moderation in the two projects can be compared, how moderators relate to the overall community of each project, and which of the roles they take on might be especially relevant for citizen science projects. The key aspects and findings of the research are presented in an article format in the second part of this thesis. The article includes the main findings and a detailed description on the theoretical framework used to discuss results and the methods used to gather, select and analyze the data. Additional information is presented in the first and the third part of the thesis. In the first part, the background of the research is explained in more detail, including a literature review on forum moderation as well as a discussion on conceptualizing citizen science projects as communities. The research design is explained and justified in more detail, followed by a summary of the findings and additional observations on possible categorizations of moderator posts. Following the article, the third part of the thesis presents further discussions about the relation of citizen science discussion forums to other discussion forums and related concepts that might be used to contextualize the findings, as well as discussions about aspects of learning in relation to discussion forums in citizen science. Some considerations on challenges, limitations and further research opportunities are presented, followed by a general summary and conclusion of the entire thesis.

## Overview and research idea

There is a growing number of projects that engage citizens in the process of scientific research. In many of these citizen science projects, information technology is used for communication, data collection and processing or collecting results; however, the forms and extent of technology use vary greatly between different projects (Wiggins & Crowston, 2011). The projects exist for wide range of topics such as astronomy, ecology, or mathematics (Franzoni & Sauermann, 2014) and often involve volunteers in the collection and analysis of data (Rotman et al., 2012). Unlike other citizen science projects, ‘virtual’ citizen science projects do not include physical elements, conducting all parts of the project with the help of information technology and connecting participants and their contributions in online platforms (Wiggins & Crowston, 2011). Therefore, exploring the interactions in these platforms, including the roles of moderators in discussion forums, seems particularly relevant for ‘virtual’ citizen science projects.

The issue of moderation in online discussion forums has been studied in different contexts. There seems to be a contrast in existing studies between discussion forums used in formal education and more publicly accessible forums. While moderation in an educational context is primarily concerned with aspects of learning and engagement, moderation in other forums focuses more on organizing contributions and excluding inappropriate content. In both contexts, moderators’ behavior is seen as a way to influence the community of participants and set the general tone for the interaction. Some authors offer categorizations of moderation styles or moderator roles derived from studied samples of discussion forums. However, there seems to be no common framework or categorization of moderator actions across different contexts.

Discussion forums of citizen science projects tend to be publicly accessible and open to contributions from a fluctuating body of participants. At the same time, many citizen science projects also state learning about the issue and the scientific process as benefit or even goal for participants. This indicates that citizen science discussion forums might lie somewhere in between the contrasting types of forums found in the literature, making them a very interesting context for analyzing forum moderation. Although there are remarks about moderation in the literature on citizen science, there seem to be no studies that consider moderation in citizen science online discussions specifically. Even though research on citizen science is emerging only recently, the lack of information on moderation seems surprising. After all, it has been suggested that moderation can impact the learning experience of participants as well as the formation of a community by organizing information and setting the overall tone for the discussion environment (Frith, 2014). Citizen science communities often involve many different participants and stakeholders, making it particularly interesting to see how moderation might affect the individual participants and the overall communities.

This thesis aims to describe moderation in the discussion forums of two ‘virtual’ citizen science projects. It is investigated how moderation of discussion forums can impact the communities of citizen science projects and how this can be compared to the two contrasting types of forums found in the literature, formal education discussion forums and publicly accessible discussion forums. The two discussion forums analyzed are the „Talk” section of the Galaxy Zoo project and the „Forum” section of the Foldit project. This allows for a comparison between moderator posts from the large body of data in the very active discussion in Galaxy Zoo and moderator posts in the much smaller and less active discussion forum in Foldit. A selection of moderator posts from each forum was analyzed and

compared to existing categorizations of moderator actions and moderator roles. The study was guided by the following research questions: (1) Which activities do moderators show in the two different forums? (2) How does this relate to existing categorizations of moderator behavior? (3) Which moderator roles are particularly relevant in citizen science projects?

It was found that the moderators' activities differ between the two datasets analyzed and also between individual moderators in the same forum. Despite these unique characteristics, moderation shows some common trends in the two citizen science projects. It seems that four common moderator activities are especially relevant for citizen science projects: Firstly, moderators can help to connect different stakeholders of the projects, especially volunteers and researchers. Secondly, they can collect and organize users' contributions in the discussion forum and information from other resources creating an overview of knowledge resulting from and related to the project. Thirdly, users' satisfaction and motivation might be influenced by moderators who recognize and praise the volunteers' contributions to the community, both in the project task and the discussion forums. Finally, moderators also seem to offer support for participants seeking to learn more about the science behind the task or other aspects of the project – to fulfill this role they do not need to be trained experts in the field but instead possess a vast range of knowledge about different project-related issues and information resources.



# Background

The background for this study includes both citizen science and forum moderation. There seems to be a lack of studies dealing explicitly with forum moderation in citizen science projects. Therefore this thesis will first consider the context of citizen science and then present a systematic literature review on the issue of forum moderation in a variety of different contexts. While many citizen science projects make use of online forums and some studies also remark on the moderation of those, no studies were found that consider the issue of forum moderation in citizen science projects as a main aspect. Consequently, a systematic literature review was conducted to find out in which contexts forum moderation has been studied so far. It will be presented in the second part of this chapter. The first part of this chapter introduces the background of citizen science, and the possibilities and challenges of conceptualizing citizen science projects as communities.

## Citizen science

Being a recently emerging field of study with a range of different objectives and backgrounds, it is no surprise that the term „citizen science” has a number of different definitions and ongoing discussions about ontological and epistemological questions.

One of the most broad definitions of citizen science is given by Edwards (2014), stating that it „involves members of the public contributing to scientific endeavours” (p. 132). Other definitions include the possible roles or actions of participants more explicitly, describing citizen science as „partnerships initiated by scientists that involve nonscientists in data collection” (Jordan, Gray, Howe, Brooks, & Ehrenfeld, 2011) or „partnerships between scientists and non-scientists in which authentic data are collected, shared, and analyzed” (Jordan, Ballard, & Phillips, 2012). Similarly, Silvertown (2009) defines a citizen scientist as „a volunteer who collects and/or processes data as a part of a scientific enquiry.” However, Silvertown also remarks upon the different notions of „science for the people” and „science by the people”, a distinction that can lead to a very different definition of citizen science: „science which relates in reflexive ways to the concerns, interests and activities of citizens as they go about their everyday business” (Jenkins, 1999). This definition is also taken back up by Roth and Lee (2004). Most of the projects described as citizen science however, do lean more towards the first set of definitions, describing citizen science as a participation of citizens in large-scale scientific projects, dealing mostly with data collection or data processing. Franzoni and Sauermann (2014) add that different terms are used to describe the phenomenon, for example crowd science, citizen science, networked science, or massively collaborative science. Furthermore, the terms used to describe contributors are discussed, noting that the term „citizen” might aim to describe a diverse community of contributors, trying to avoid stereotypical, limiting or negative connotations of other terms such as volunteer, amateur, or non-scientist (Edwards, 2014). Furthermore, there are different approaches to the categorization of citizen science projects. This includes for example a distinction by the influence of volunteers; differentiating between contributory, collaborative, and co-created citizen science projects (Rotman et al., 2012). Another typology considers more organizational properties like project goals or the importance of the physical environment to distinguish five types of citizen science projects: Action, Conservation, Investigation, Virtual, and Education (Wiggins & Crowston, 2011).

Although the use of volunteers for monitoring projects has a long history in some academic disciplines (e.g. ornithology), the term „citizen science” is only emerging more recently (Silvertown, 2009). The range of projects involving the public in scientific research is increasing (Silvertown, 2009) and

moving into more academic disciplines (Bonney et al., 2009), leading to a large number and diverse range of projects (Edwards, 2014; Franzoni & Sauermann, 2014). The increasing involvement of the public can be attributed to several factors. Silvertown (2009) lists the evolution of technological tools, the recognition of the public as a valuable resource, and the increasing need for justifying research and reaching out to the public as factors contributing to the rise of citizen science. Wiggins and Crowston (2010) also point out that „ubiquitous computing makes broad participation by the public in scientific work a realistic research strategy for an increased variety of scientific research problems” (p.151). Furthermore, citizen science is gaining importance as it might give learners and non-experts an opportunity to participate in authentic scientific practices, providing them with valuable learning opportunities (Raddick, Bracey, & Gay, 2010) and opening up possibilities for lifelong learning and participation in the community (Edwards, 2014; Roth & Lee, 2004). Citizen science has been described as a particularly important tool for ecological studies, enabling large-scale observations and monitoring, providing opportunities to partner with relevant stakeholders in local communities, and influencing public engagement and learning both on the specific topic addressed and on science in general (Dickinson et al., 2012). It has also been argued that citizen science might in combination with science education and environmental education help to reform school curricula to address sustainability issues (Wals, Brody, Dillon, & Stevenson, 2014).

The recognition of citizen science as a learning opportunity leads to an increase of projects aiming to contribute to both scientific and educational objectives (e.g. Bonney et al., 2009). Research on citizen science has explored different ways of evaluating learning outcomes and impact on participants’ attitudes, both in relation to the specific task as well as in relation to science more generally. This includes studies by Brossard, Lewenstein, and Bonney (2005), Crall et al. (2012), Jordan, Gray, Howe, Brooks, and Ehrenfeld (2011), and Trumbull, Bonney, Bascom, and Cabral (2000). While impacts on knowledge and attitudes have been found, the findings differ as to how generalizable the learning and attitude changes are, indicating that the relation between participation and learning is complex and not necessarily direct. Additionally, Jordan, Ballard, and Phillips (2012) point out that the evaluation of community-level outcomes such as social capital, trust, or resilience should also be considered and evaluated in addition to individual learning.

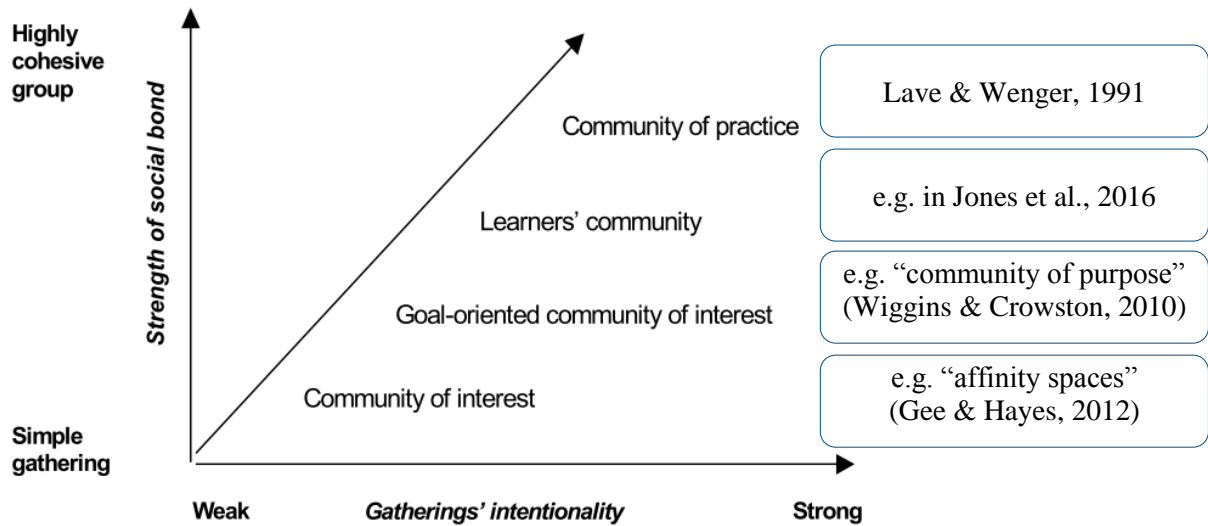
## Can citizen science projects be conceptualized as communities?

Different frameworks have been used and developed to describe citizen science projects as communities. The broad range of possible projects leads to a variety of possible interactions and stakeholders. Local small projects aim to involve existing communities in relevant research, rather than shaping new communities. However, large-scale projects and virtual projects like Galaxy Zoo and Foldit bring together a range of different stakeholders from various backgrounds into a collaborative project. For these projects, theories and frameworks from online communities based on a common interest might be applicable. A range of theoretical frameworks offer descriptions of online communities, many of them explicitly contrasting the concept of communities of practice.

The concept of communities of practice including legitimate peripheral participation (Lave & Wenger, 1991) seems a promising approach to analyze virtual communities at a first glance. Studies can be found that describe citizen science communities as communities of practice with an opportunity for legitimate peripheral participation (Jackson, Østerlund, Crowston, Mugar, & Hassman, 2015; Rotman et al., 2012). However, both studies complement their theoretical frameworks with additional concepts theories. A different well-known theoretical framework was used for a locally limited citizen science

project by Roth and Lee (2004): The authors use cultural-historical activity theory (CHAT) to describe citizen science as an activity system that can be represented by the relation between subject and object and mediated by different factors. Wiggins and Crowston (2010) state that core-periphery models generally might present good descriptions of citizen science projects but point out that there is a very high difference in formal status between the scientists as core members of the research project and citizens or volunteers as peripheral participants. The authors argue that in the case of citizen science, the common element between participants is not so much a shared practice but rather a shared output. They develop their own model of citizen science organizations, based on an input-mediator-output-input (IMOI) model that (similar to CHAT) includes factors that can mediate the elements of the model. In their work, Wiggins and Crowston draw on ideas from other massive virtual collaboration projects like social networking, or open source software development. In a later publication, they note that similar characteristics to open communities and peer productions make research on those related phenomena a useful background for their conceptualizations on citizen science (Wiggins & Crowston, 2011). In their work on citizen science virtual organizations, Wiggins and Crowston (2010) remark that „crowdsourcing” can point to different forms of massive collaboration and is not a very well-defined term. A distinction between the different but related concepts is offered by Franzoni and Sauermann (2014) who differentiate between four types of projects, depending on the openness of participation and the disclosure of intermediate outputs: traditional science, crowdsourcing, traditional science with disclosure, and crowd science. Similar to Wiggins and Crowston, they compare the citizen science to open source software development and open innovation and argue that practices of organizing information are similar.

Considering the many similarities between different crowdsourcing approaches, theoretical frameworks from those approaches might also be useful to consider in the citizen science setting. Models that consider the fluctuating membership and present themselves as alternatives to the idea of communities of practice include the idea of affinity spaces (Gee & Hayes, 2012) or mycorrhizae („knot-working”) (Engeström, 2007). However, in these models, the distribution of tasks and roles is much more informal than in most citizen science projects. Henri and Pudelko (2003) might offer help for choosing a relevant and fitting model for citizen science communities: the authors present different types of communities based on their social bond and intentionality (see figure 1).



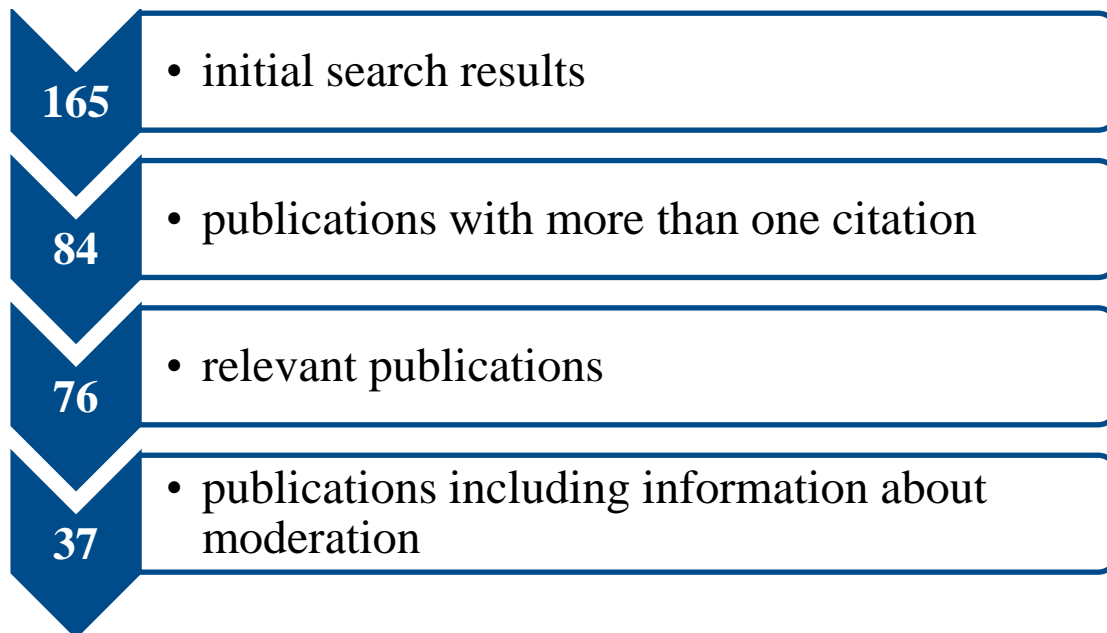
**Figure 1: Different forms of communities (Henri & Pudelko, 2003) and examples of theoretical frameworks that can describe them**

This model can not only be used to describe the different types of communities, but also offers itself as a way of distinguishing the different theoretical approaches that might fit them. Using this distinction to describe the different theories, the models by Gee and Engeström would resemble „communities of interest”, characterized by weak social bonds and low intentionality. Communities of practice and learners’ communities on the other hand, have a much stronger intentionality and social bond that could make more formal models like Wenger’s peripheral participation applicable. In a study on a MOOC (massive open online course), researchers noted that the two concepts of affinity spaces and communities of practice might both present valuable insights about the community (Jones, Stephens, Branch-Mueller, & de Groot, 2016). If the MOOC is considered as a learners’ community, it would indeed fit in between the two types of communities that the two concepts describe. Citizen science communities on the other hand could be described as goal-oriented communities of interest, having a stronger intentionality than simply communities of interest, but much weaker social bonds than formal learners’ communities or communities of practice. This indicates that a model for a goal-oriented community of interest, like Wiggins and Crowston’s (2010) conceptualization of a „community of purpose” might describe the unique characteristics of citizen science projects very well, accounting for the differences (e.g. formality, top-down approach) and similarities (e.g. fluctuating membership, challenges of communication and collaboration) to related practices. Therefore, the framework of virtual organization by Wiggins and Crowston was selected to discuss aspects of forum moderation in this study. The model is explained in more detail in the article section on the theoretical framework.

## Literature Review on Forum Moderation

Since there seemed to be no discussions of forum moderation in citizen science projects specifically, a systematic literature review was conducted to cover aspects of forum moderation in different contexts. In particular, it was tried to discover how and in which contexts moderation of discussion forums has been studied, which roles participants can take on, and which different types of moderation occur. The review showed that a majority of the studies on the topic of moderation in online discussion forums was conducted in the context of education, in most cases higher education. Most of the studies used an analysis of forum data, participant interviews or a combination of methods for collecting data. The

literature suggests different roles that participants can take on (and are assigned to in an educational context). Different aspects and types of moderation are discussed, with a notable difference on purpose between educational discussion forums aiming to facilitate learning and more general discussion forums aiming to control inappropriate content.



**Figure 2: Overview of the literature selection process**

The literature review on forum moderation was conducted with a systematic approach. The database Scopus was used to discover relevant literature. A total number of 165 publications were obtained in the database search, and subsequently narrowed down to 39 publications that were analyzed in more detail. Figure 2 shows an overview of the selection process of the publications. The first search (limited to English papers, published up to and including 2015) resulted in 165 publications on the topic of moderation in online discussion forums. For an initial analysis, the tools provided by the Scopus database were used to obtain information about the entire dataset. For a more detailed analysis, the 84 publications cited more than one time were screened more carefully. 7 search results were manually excluded since they did not concern online discussion forums, leaving 77 relevant articles. For these articles, the abstracts were examined to determine what types of forums the research was dealing with and if the publication was about discussion forums in general, discussion forums as well as moderation, or the topic of moderation in particular. 37 publications were identified that seemed to consider aspects of moderation (39 dealt with other or more general questions on discussion forums). These articles were then analyzed in more detail (using MS Excel), considering especially the forum studied and the research methods, theories and concepts used, as well as the purpose and findings of the studies. Common themes were identified and will be discussed in the following sections. The detailed analysis of the 37 papers was used to gather information about the context of the studies, possible user roles, and types of moderation.

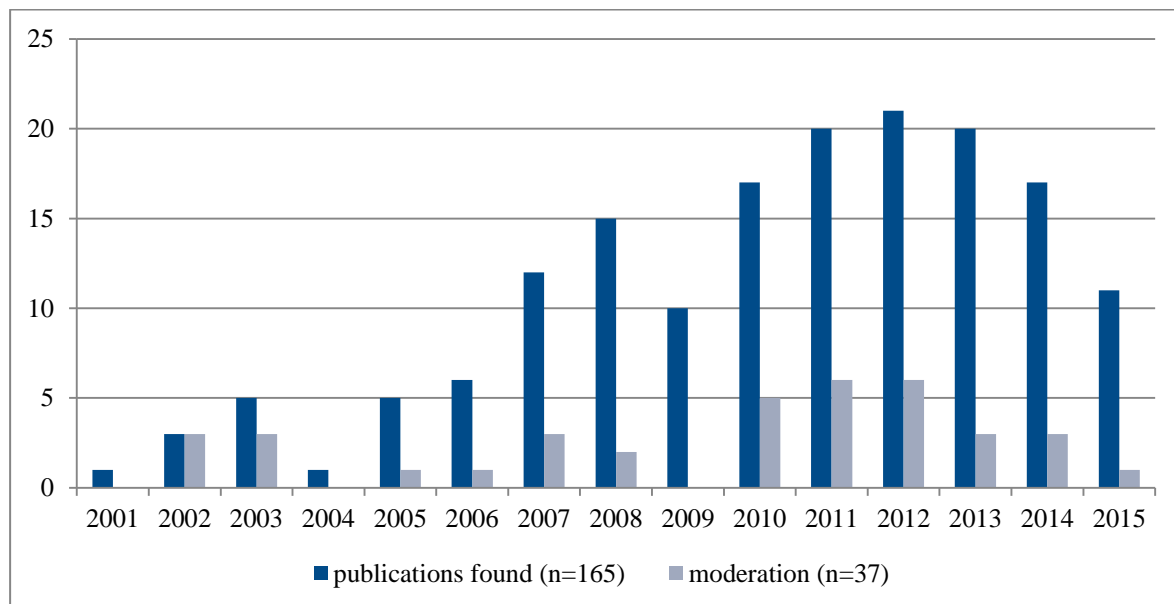
Constructing search terms to find relevant publications proved to be challenging, since many different terms are used to refer to online discussion forums (including online forum, discussion forum, only community, message board, asynchronous discussion, internet forum). Furthermore, the terms „moderation” or „discussion” can be used in different contexts, leading to the inclusion of many

irrelevant search results. By conducting and comparing several searches and looking at the keywords of initially identified papers, the following search query was established:

TITLE-ABS-KEY (moderat\*) AND (KEY (discussion) OR KEY (forum)) AND (TITLE-ABS-KEY (online) OR TITLE-ABS-KEY (forum) OR TITLE-ABS-KEY (communit\*))

Additional terms for online forums were added to the search as well but did not seem to improve the search results.

After obtaining the initial dataset of 165 publications, the analytical tools provided by the Scopus database were used to get a first overview of the documents. Figure 3 shows the distribution of publications by year, indicating the distribution for the initial dataset as well as the final selection of 37 documents. The first relevant documents identified in the search were published in 2001; the first year with more than 10 publications was 2007. There was a peak of 21 publications in 2012 (with 20 publications in 2011 and 2013 respectively), and a slight decline in 2014 and 2015.



**Figure 3: publications on forum moderation by year**

The sources of publications were mostly from the fields of computer science or education, with the most documents found for Computers in Human Behavior (9), followed by Computers and Education (6) and Educational Technology and Society (4). No authors or affiliated universities stood out much from the dataset, with the maximum of publications for one author or institution being only 4 in both cases. Out of 36 identified countries, the most publications were recorded for the USA (54), followed by the UK (54), Australia (16) and Germany (8). Slightly more than half of the identified documents were listed as articles (90 articles, 54.5%), 62 (37.6%) as conference papers. Of the remaining documents, 7 were listed as reviews, 3 as book chapters and 3 as articles in press.

Many different subject areas were listed for the dataset, which suggests that online discussion forums are used in different areas and for different purposes. The majority of the documents were from the subject area of computer science (92 documents, 55.8%), followed by social sciences (88 documents, 53.3%). Much smaller numbers were reported for other subject areas, with engineering (25 documents,

15.2%), medicine (17 documents, 10.3%) and psychology (16 documents, 9.7%) making up the biggest groups. 13 other subject areas were also identified. Adding up even the biggest two groups already gives more than 100% of the documents, suggesting that many of them fall into more than one subject area, i.e. computer science and the respective area of specific forums studied.

Skimming the titles and abstracts of the papers showed that many of them seemed to stem from an educational context and that moderation was still only a passing remark in many of the publications. Therefore, the 76 publications that were cited more than once and relevant to the topic were analyzed by titles and abstracts in order to identify in which context forums were analyzed and if the paper dealt with forums in general, forums as well as moderation, or moderation in particular. An overview of this first analysis can be seen in Table 1.

number of documents		topic of the paper				
		forum	forums and moderation	moderation	TOTAL	%
type of forum	Professional communities	3			3	4%
	consumers	2	1		3	4%
	political debate	7	3	1	11	14%
	education	21	17	8	46	60%
	medical advice	3			3	4%
	other	4	3	3	10	13%
	sports			1	1	1%
	<b>TOTAL</b>	<b>40</b>	<b>24</b>	<b>13</b>	<b>77</b>	
	%	52%	31%	17%		

**Table 1: publications by topic and type of forum**

This first classification of the papers revealed that more than half (46 documents) of the publications were from an educational context. The only other context with a significant number of documents was that of news and political debate. Other identified contexts contain only few documents each. Of the 77 documents, only 13 dealt explicitly with the topic of moderation; 24 also included aspects of moderation even if it was not the main aspect of the paper. These 39 documents together were analyzed in more detail in order to find out how and in which contexts moderation has been studied, which roles participants can take on, and which different types of moderation occur.

### **Studying moderation in online discussion forums**

The variety of possible search terms and the many different subject areas identified in the preliminary assessment of the full dataset (n=165) seem to suggest that discussion forums have been studied in different contexts and different academic disciplines. However, at a closer look, it turns out that the forums studied and the methodology used concentrates mostly on the analysis of forum data in an educational context. Most of the studies rely on the qualitative analysis of forum content, interviews with participants, or a combination of these two formats for the analysis of data.

The first classification of papers revealed that many of them seemed to stem from an educational context. Considering the 37 papers related to moderation, an overwhelming majority (25 papers) considers educational forums. Out of these, 21 focus on higher education (3 studies were conducted in

high schools, 1 in professional development). 6 documents deal with general news or political discussion boards, 3 are only theoretical works considering no particular context, and 3 consider other contexts (medical, technical, sports).

The variety of subject areas found in the first overview of the papers can be explained by the variety of subjects taught with the help of online forums in distance education or blended learning, including for example medical education (De Wever, Van Winckel, & Valcke, 2008; Thomas, 2013), engineering education (e.g. Danchak & Kenyon, 2002) or instructional design education (De Wever, Keer, Schellens, & Valcke, 2010). Although the studies focus on different subjects taught as well as different aspects related to moderation, for example participation or knowledge building, the concentration on educational contexts means that learning is an important factor in most of the discussions. The forum software used also reflects the educational environment, in some cases being the discussion forums provided by learning management systems (e.g. Moodle in Hsieh & Tsai, 2012; Jyothi, McAvinia, & Keating, 2012).

The focus on educational contexts also results in many studies that consider only a limited time frame, i.e. the duration of a course or learning module, a limited number of participants and a very controlled discussion environment. In some cases, a co-located set-up was used which limited the interactions that were found in the forum discussions, since presumably they were done offline (Asterhan, 2011).

Most of the studies included in the detailed review (n=37) use some form of analysis of forum data, interviews with forum participants, or a combination of methods. Content analysis of the discussion forums seems to be the most used approach, followed by interviews with forum participants or forum moderators. Many studies use several methods in order to triangulate data. Only one study uses the description „online ethnography” (Dong, 2012), and one study uses social network analysis (Xie, Yu, & Bradshaw, 2014). Three studies use grounded theory for the analysis of data (Frith, 2014; Vlachopoulos & Cowan, 2010a, 2010b). Several studies use comparisons and experimental or quasi-experimental designs. This includes a comparison between synchronous and asynchronous communication (Chou, 2002) and comparisons between moderated and unmoderated discussions (Herman, 2010; Hsieh & Tsai, 2012). More technically or theoretically oriented studies also propose frameworks or build software solutions, testing them on data from discussion forums (e.g. Cerulo & Distant, 2013). One of the publications is a literature review on forum discussions in medical education (Thomas, 2013), another uses an extensive review of literature to develop an analysis method for moderators in education (Brace-Govan, 2003).

It is notable that learning or knowledge building as well as participation or engagement are factors considered by almost all of the studies from educational contexts, whereas other factors like motivation, control of the contributions or technical support do not seem to play an important role in the discussions in that context.

### **Roles of users in discussion forums**

Participants in online discussion forums can take on many different roles, including, but not limited to that of a moderator. Several studies are concerned with different roles that students can take in educational online discussion forums (and how this affects their learning and participation). However, it is pointed out that „Researchers assign and study different collections of roles, and even when similar role names are used (e.g., moderator), specific responsibilities vary” and that the



responsibilities of roles should be noted carefully in studying them (Wise, Saghafian, & Padmanabhan, 2012).

De Wever, Keer, Schellens, & Valcke (2010) assigned students to five different roles (starter, summariser, moderator, theoretician, source searcher), suggesting that the roles might help knowledge building if introduced at the beginning of the discussion. In another study, a higher level of knowledge construction occurred when the role of moderator and of a developer of alternative ideas were both given to students (De Wever et al., 2008). Similarly, (Xie et al., 2014) found that students assigned to moderating were more engaged. Another study identified seven common functions learners are asked to perform [motivate others to contribute, give direction to the conversation, provide new ideas, use theory to ground the discussion, bring in (relevant external) sources, respond to previous comments, and summarize existing contributions] and connects them to six role descriptions (Wise et al., 2012).

All in all, the research that discusses roles of forum participants tends to focus on forums in educational contexts only. Researchers are especially concerned with the question how assigning roles to students impacts their participation and their learning, while roles that participants might take out of their own motivation are not discussed. Only one study on moderation in a sports forum considers the motivations of volunteer moderators (Alonso, O', & Shea, 2012).

### **Styles of moderation**

Different styles of moderation are discussed in the studies considered for the literature review. Most notably, there is a distinction in the purpose of moderation. Furthermore, the behavior and required skills of moderators are discussed. Finally, there are some studies concerned with opportunities of automatic or semi-automatic moderation.

### ***Purpose of Moderation***

The most notable difference between studies from educational backgrounds and from other contexts lies in the purpose of moderating discussions: While the studies from educational contexts focus on fostering engagement, participation and learning of the participants (e.g. Kienle & Ritterskamp, 2007), news discussions or more general discussion forums tend to employ moderation in order to control inappropriate content. The difference becomes very clear in comparing the problems and proposed solutions stated in the different papers. In an educational context, the role of moderating is still closely connected to teachers and moderation is aimed at influencing the participants: „This tool can help teachers or moderators to intervene in the discussion if necessary, changing the participants' focus and activity.” (Jyothi et al., 2012) A focus of moderation lies on assessing the students' contributions and deciding if and how to intervene (Brace-Govan, 2003). In more general discussion forums, the focus of moderation lies on inappropriate content: „This paper addresses the problem of moderating [online discussion sites], and it presents a novel technique for automatically identifying contributions not complying with a site's terms of use.” (Delort, Arunasalam, & Paris, 2011)

It is also suggested that analytic tools for moderation can serve a double purpose, allowing the analysis of discussions not only for moderation but also for research: „Furthermore, the visualizations generated can support deeper analysis, including qualitative and quantitative research, into student learning in online discussion forums.” (Jyothi et al., 2012) All in all, many of the studies confirm that moderating the participants in some form leads to a better (learning) outcome; a study on a counseling intervention for career development by Herman (2010), finding that „professional moderation resulted in better outcomes on several variables and greater overall satisfaction with the intervention.”

### ***Moderator Behavior***

Many of the examined publications discuss the behavior of moderators. Again, there are strong connections to participation and learning and studies try to examine which moderating strategies have the best impact on students' performances.

Vlachopoulos & Cowan (2010a) identify six approaches of moderation between „tutoring, 'managing' and 'facilitating'" student discussions online: One track mind (focusing on final goals), Top of the list (prioritizing), Going the second mile (encouraging learning beyond demand), Critical friend (collegial relationship), balancing priorities (multi-tasking), rescuing (avoiding disaster).

Asterhan (2011) differentiates „a scaffolding, an orchestrating, an authoritative, an observing and a participative style" of moderation and identifies five moderator action categories: pedagogical scaffolding, interaction support, managerial support, involved discussant, moving forward.

Furthermore, she notes that social and technical support that might occur in other settings was not visible in the dataset due to its set-up as a co-located discussion with support given directly by the teacher on-site.

Guldberg & Pilkington (2007) note that the nature of questions posed also influence the discussion and that the key to successful moderation might in fact lie with the preparatory work before leading online discussions, rather than the response to participants. Other studies also remark on the style of moderation and moderators' behaviors, pointing out the challenges in moderating and choosing a style of moderation: „The extent to which the e-moderators were clear or not about their roles in online discussions, and were ready to adhere to them, thus directly influenced the e-moderation style which they adopted, as well as their purposes for intervening" (Vlachopoulos & Cowan, 2010b)

Xie et al. (2014) note that successful moderation might require training and (Danchak & Kenyon (2002) state that moderation might be especially challenging for educators: „Instructors are used to being the expert and it is difficult for them to adjust to the role of moderator." Frith (2014) suggests that technical communicators might already possess the skills needed to manage an online community, especially in the context of technical help forums. Additionally, a finding from political debates suggests that effective moderation styles might also depend on group structures. „The findings suggest that different group compositions and purposes may lead to different moderation as well as participation behaviors, which result in the different quality of online political discussions as we have seen in the eight cases." (Zhang, Cao, & Tran, 2013)

### ***Automated Moderation***

It is mostly news and general discussion sites that are considered for discussions on automatic moderation, due to the large amount of content generated on those sites: „As the volume of user-generated content (UGC) increases, a solitary trusted moderator cannot single-handedly deal with the problem of identifying bad content" (Ghosh, Kale, & McAfee, 2011). Algorithms are suggested and tested in order to identify inappropriate content automatically or semi-automatically (Delort et al., 2011). Furthermore, automated moderation also includes the rating or scoring of individual contributions for recommendations as a way of dealing with large amounts of information (Arnt & Zilberstein, 2003; Wang, Li, & Chen, 2010). While most of these approaches refer to non-educational discussion forums, there also is one study suggesting „topic-driven semi-automatic reorganization" of contribution for a large university discussion board: „Discussion forums represent one of the main asynchronous communication means offered by any learning management system and discussions

taken place and stored in them represent a source of information for learners accessing the forum afterwards. Their effectiveness as information sources, i.e., the capability to satisfy users information needs, depends on their information richness first, but also on how discussions are organized and effectively moderated.” (Cerulo & Distanto, 2013) The discussions about automated moderation make the different focus of studies from educational and other contexts more obvious. They also reveal that moderation might refer to both the identification of bad content as well as the (re-)organization of content in order to help users to find the needed information.

To sum up, the first analysis of the dataset suggested that discussion forums can be used for a large number of contexts and subject areas and that most of the research concerns more than one subject area (in most cases computer science and the topic of the specific forum analyzed). However, the detailed analysis of the documents shows that the study of moderation is very focused on educational contexts, especially within higher education. Many aspects studied are particular to educational contexts, while more general discussion forums deal with different problems, such as the necessity of automated moderation to fight inappropriate content. Many of the studies presented analyze only short and very limited interactions of students in specific courses who were asked to discuss specific tasks. The discussions about user roles in online discussions show that many different categorizations of roles are possible. This means that conclusions about „moderator” or other roles should be considered within the context of the specific study and discussion forum and are not necessarily generalizable. Even though some common themes can be identified, it should also be noted that the research mostly dealt with assigning roles to students and did not consider roles that discussants took on by their own account. The styles of moderation identified are in close relation to the purpose of moderation and the behavior of moderators. Several authors suggest that even within the limited settings of educational online discussion, careful assignment of roles is important and that training moderators might be important. It is noteworthy that a distinction is made between the role of a teacher and that of a moderator (Danchak & Kenyon, 2002) and that one study finds better results in assigning students (not instructors) to the role of moderator if another student is appointed as a developer of alternative ideas (De Wever et al., 2008). A theme of automated moderation emerged from the literature, showing that this topic is mostly a concern for big and general discussion sites but also indicating that it might become important in the field of education as well, especially when it is considered not for the removal of bad content but for the organization of information to help users (Cerulo & Distanto, 2013).

The variety of terms used to describe forum moderation and the many different contexts make it challenging to identify relevant literature. However, the literature retrieved from the Scopus database for this review shows that the topic of moderation in online discussion forums is very concentrated on the context of (higher) education. Some of the literature from other contexts shows that the purpose and style of moderation might differ depending on the discussion forum considered. It seems that more research in different contexts of forum use might be valuable. In particular, the analysis of online discussions within less controlled environments and over a longer period of time seems to be missing from the literature. Additionally, the topic of moderation seems to be considered only in very specific contexts, while the few reviews obtained in the dataset deal with discussion forums in general, mentioning moderation but not considering it in depth.

While citizen science projects also seek to engage participants and foster learning similar to educational settings, discussion forums in those contexts operate very differently from those in controlled formal educational settings. The roles of participants cannot be assigned or assessed

formally, and there is a big fluctuation of participants. Nevertheless, the relation to learning is present in citizen science and might still require a different approach than a filtering of inappropriate content as it is emphasized by news discussion forums. While transferring some of findings on moderation to the context of citizen science might be appropriate, the different settings should be considered very carefully.

## Research Design

A full description of the methods used to collect the data from the forums, select a subset of posts for analysis, and categorize the posts, can be found in the article's chapter on methods. In this section, some additional details about the research design are explained.

After obtaining the datasets from the two discussion forums, the more challenging task was to narrow down the large datasets to a selection that made analysis within the time frame of a masters' thesis feasible. Additional methods like social network analysis or concordance analysis were considered to find relevant actors or posts for analysis. However, while these methods might have given interesting results and possibly allowed work with a bigger set of data, they present additional concerns and challenges more suitable for separate research projects than for including them in the frame of this thesis. Therefore, the more simplistic selections by time (Foldit) and by subforum (Galaxy Zoo) were chosen to be able to concentrate on the analysis of a smaller selection of moderator posts.

From the literature review on forum moderation, it became clear that a wide range of methods is used to study forum moderation. Common methods include content analysis and interviews, sometimes in combination, and in some cases adding forms of structural analysis. Therefore, a similar approach was chosen for this study: selected posts were analyzed with regards to their content, while additional observations on the structure of the data were also noted. Some general information about the forums and the different datasets was collected and presented to give readers an impression of their differences and similarities. For the selected datasets of moderator posts (91 in Foldit, 104 in Galaxy Zoo), each post was considered carefully and categorized with regard to moderator actions as identified by Asterhan (2011) and moderator roles observed in Frith's (2014) interview study on moderators in technical help forums. The two possible categorizations were chosen from different backgrounds (educational, closed forum and technical, open forum) to account for aspects from both backgrounds that might be present in the citizen science discussion forums. Additional notes taken during the categorization as well as comparing different posts to each other allowed for further scrutiny of the data to find fitting examples for different categories and to figure out themes for discussion that emerged from the data .

### Ethical concerns of the data collection

Internet based research poses special challenges and ethical concerns about data collection: Although discussion spaces are publicly accessible, users often reveal private information and obtaining permission for the use of the data and ensuring anonymity might be necessary, but informing participants about the study might alter their behaviour (Convery & Cox, 2012). In this study, the focus on moderators of well-known citizen science project means that the moderators are easily identifiable, even if their user names are anonymized. However, in addition to being publicly accessible and visible even without logging in to the platform, data from the projects and the discussion forums has been used in a variety of studies. It is assumed that most members and particularly the moderators are well aware that their messages and contributions are in fact publicly available and might be used in research. Furthermore, the data collected contains no sensitive information and the moderators are seen only in their function as moderators (or, to some extent, as participants of the projects); no personal information was revealed or synthesized in this study. Since the details of the project are important for the context of the research, it was not feasible to leave out the forum names to protect moderators' anonymities like Frith (2014) proceeded in his study on

moderation. To prevent an identification of moderators on topics that might cause concern to them, posts that seemed too controversial or personal were not cited as examples but left out or summarized in different words. Unlike Frith's interviews, the data for this study is publicly available and unlikely to reveal controversial or personal aspects about the moderators. Additional methods, like collecting data from interviews or a social network analysis describing the forum participants, might pose ethical concerns in relation to revealing new, previously unknown or private data; however, although these methods were at first considered for the study they were not carried out due to time constraints.

## Findings

The literature review on forum moderation revealed that there is a different perspective on moderation between educational and other types of forums. It was suspected that citizen science might display some characteristics of both, since the projects include both educational aspects as well as open participation. Indeed, almost all categories of moderation aspects identified in an educational setting by Asterhan (2011) and in technical help forums by Frith (2014) can be found in the selected datasets from Galaxy Zoo and Foldit. The focus of the moderator activities varies slightly between the two forums and between individual moderators.

The findings of the study with detailed presentations on the forum contexts as well as descriptions and examples of moderator actions following Asterhan's distinctions can be found in the article section of this thesis. Additionally, numerical data on the forum activity and moderator involvement in general can be found in Appendix 1. In summary, the numerical data shows that the two discussion forums follow a pattern typical for general discussion forums: a small number of very active core users (including the moderators) with many more users that contribute only infrequently and sparsely. However, similar to Asterhan's observation in an educational context, the moderators' posts only make up a small percentage of the overall discussion (the moderators seem to become the most active members only over time, due to the fluctuation of other members). In addition to the findings from the article, the following sections present more observations on different categorizations.

### Long posts with several actions

Some of the moderator posts fall into more than one category or include several aspects of moderation. In these cases, considering the details of the post reveals that they might include several moderator actions as categorized by Asterhan (2011). The following complete post from the Foldit forum can serve as an example:

*Hello there! While we cannot speak to specifics about your group members, here are some general tips to help you get everyone back on track- 1. Every client being used by students must be configured to have chat disabled (there is no „overriding“ teacher option, although we've certainly discussed how handy that may be in the future). 2. Be sure to restart the client to make sure these changes take effect! This will allow students to participate in the non-public channels. While we do welcome student groups in global chat, we understand that it can often be a distraction to lessons, or a possible temptation to violate our well meaning community rules (enforced ably by myself and our volunteer player moderators) and hope these tips will help get your group into an environment that works best for you and your learning needs. If specific group members of yours have issues with accessing group only chat after following these steps please feel free to drop me a private message. :)*

This post is taken from a thread started by a middle school educator who is experiencing difficulties with students in the group chat function (that should have been made inaccessible for the students in the school setting). The moderator writes out a long and friendly answer for the educator. The overall message is clearly focused on technical support, trying to help with trouble-shooting and getting the students „back on track“. However, it can also be argued that this post shows some form of pedagogical support (or at least technical support for a stakeholder especially concerned with pedagogy): this and other posts by and for educators on the forum reveal stakeholders not immediately

visible in the overall setup of the project, students and educators. It seems that the game is used by them for educational purposes (although it has a steep learning curve and is not designed as an educational game). The moderator supports this use of the game by offering information and support for the educator. In some sense, this post also provides interaction support, giving the educator the opportunity for further private messages and indirectly guiding the students' (future) interactions with the project and communication channels. Looking at the wording in this post can give some further insights: the moderator uses the plural form („we”) throughout the text, speaking on behalf of the Foldit team or even the entire community. Finally, Asterhan's (2011) observation that social support might be found in subtle forms and expressions, can also be seen in this example: the moderator begins the message with a greeting, concludes with a smiley face, is very careful to point out that only „general tips” can be provided, and acknowledges the need for „an environment that works best for you and your learning needs”. This aspect of moderation can also be compared to Frith's idea of moderators as tone-setters. To conclude, many posts can clearly be identified as examples for distinct categories, but some, especially longer posts, can contain different aspects of forum moderation in the context of citizen science.

## Moderator roles

Although Frith's (2014) identification of moderator roles results from interviews with moderators and not forum posts, his observations might also provide useful insights in this study. Therefore, the categorizations from Frith will be used to search for examples for the respective roles. The idea of „moderators as quality control experts” was not visible in the datasets, which is not surprising since spam posts usually are removed fast in discussion forums; it is impossible to tell from the moderator posts if they are responsible for this aspect of moderation in Galaxy Zoo and Foldit. However, examples of the other four moderator roles presented by Frith can be found in the datasets.

### **Moderators as information architects**

Among the tasks for moderators is that of shaping the overall architecture of the forums' content (Frith, 2014). In both Foldit and Galaxy Zoo, „sticky” threads provide information and support for beginners and try to answer frequently asked questions. One moderator stands out with a large number of posts, very long posts, and many contributions to sticky threads in the Science/Science subforum of Galaxy Zoo. This includes advice for learning more about astronomy („*If you want to learn more about galaxies, here are links to useful information from Talk, Forum and the SDSS website and other places [...]*”) as well as a try to build an index of the Forum: „*Index for Galaxy Zoo Talk [...] Where to find posted galaxy information in Talk, the old Forum, Scientist's Blogs, Zooniverse citizen science websites... I am starting this because I can't find things, even things that I have posted myself! [...]*” The moderator tries to bring together information from a variety of sources and create an overview for other users (and themselves). This includes both sources inside the project (especially the current forum 'Talk' and an older forum version) as well as external sources from different backgrounds.

The previous statement by the moderator hints at another aspect of 'information architecture' also mentioned by Frith: constraints within the website. The index was built out of frustration of not finding information in the way the system is set up. A variety of sources, including both a current and an archived forum, makes it difficult to locate relevant information quickly. Furthermore, the forum is set up to connect very deeply to the task: Users can discuss the objects presented in the classification procedure and the forum is an integral part of the website's software system. A button on the top right corner invites users to „Return to classifying” immediately. In Foldit, the forum is much more



separated from the task. To be able to play the game, users must first download a separate client. In the dataset for this study, forum posts for particular groups or puzzles were not included since they are not immediately accessible (they would correspond more closely to the discussion of objects in Galaxy Zoo). Furthermore, Foldit offers a range of other channels for information and communication. In addition to the forum and blog, there is a chat feature built into the game client. A wiki built by the community is used to collect information and provides an overview of different topics. And posts in the forum even mention other media like skype or google hangout that are used to communicate.

### **Moderators as tone-setters**

The aspect of „moderators as tone-setters” (Frith, 2014) is closely related to Asterhan’s (2011) moderator action of social support. However, for Frith, this characteristic is identified not in individual posts but across the entire forum, manifesting itself for example in how welcoming (or restricting) beginners’ posts are answered. Furthermore, Frith states that moderators can take a very active approach in shaping the forum, or be more reactive and step in only if necessary. In both forums studied for this thesis, greetings and emoticons are used frequently, and users are often referred to by mentioning their nicknames, especially when thanking them for contributions or answering their questions. This helps to create a polite and friendly atmosphere. However, there also is a difference between the two forums: In Galaxy Zoo, the moderators are more active in sharing resources, whereas additional information in Foldit mostly seems to be provided by other users and is acknowledged by the moderators. The longest of the 91 messages analyzed for Foldit contains several statements about the overall atmosphere and user base in the community: „[...] *Foldit attracts a unique type of player [...] and while quantity can be important, Foldit tends to retain a certain type of player that is not easily found in the „mass market“ sphere. [...] It’s of course a team effort, as we expect everyone to be pleasant, helpful, and welcoming when new people show up. :) [...]*” While new members are welcome, the moderator is aware that Foldit can only retain few highly talented and dedicated players, which corresponds to the project goals and the very competitive game. On the other hand, Galaxy Zoo is much more open to a wider range of contributions, offering both simple and complex explanations and resources. This reflects in the overall quite friendly tone of the forum and moderators’ tries to provide answers in simple terms, e.g. „*Those green thingies are optical artifacts*”. Furthermore, it seems that the moderators of Galaxy Zoo place more emphasis on acknowledging and celebrating contributions to academic publications.

### **Moderators as translators**

Frith’s (2014) study suggests that moderators might work as translators between users with different levels of experience. Being part of a global project, one Galaxy Zoo moderator even comes across literal translation issues: „*Actually, we try to keep things in English, so everybody can understand it. But feel free to send me a pm (personal message) in Dutch. Or just post your question in Dutch and we’ll take it from there ! :-)* Glad you to have you on board and happy Hunting !” However, although the moderators do answer questions across different levels of expertise, no posts were found actually showing a „translation” between different levels of expertise.

However, Frith’s idea can be interpreted differently and might be particularly relevant in the area of citizen science: moderators in citizen science projects can act as intermediaries between different stakeholders of the project and community. The following examples are from the Foldit discussion forum, similar examples might be found in other, more technical subforums of Galaxy Zoo as well. In one thread, the moderator is collecting input from users for further development: „[...] *As always, your input is important to us, and we try to maintain a strong balance between „hard core science*

goals“ *from the lab, and your ideas of „what makes Foldit fun“*. [...]” This suggests that the moderator works as a translator or intermediary between the players on the one hand, and the researchers and developers on the other hand. In another thread, the moderator reveals a close connection to the project developers: *„Our devs had two suggestions as to why this may happen [...]”*. In several more examples, the moderator offers to pass forward information to the development team. The other way around, it often is the moderators who announce (and maybe even defend) new developments to the software or the overall project. In Galaxy Zoo, the moderators are very active at pointing out the scientific publications that result from the project, which might be seen as another act of translation or mediation between stakeholders.

### **Moderators as knowledgeable non-technical experts**

Frith (2014) explains that many of the moderators interviewed in his study were not technical experts for the field of the forums they were moderating. However, they did show a broad range of skills and expertise in dealing with the technicalities of forum moderation and ‘translating’ expert knowledge. The moderators of the forums for Galaxy Zoo and Foldit might not be professional scientists, or even very active in contributing to the project task. However, they do show a broad range of knowledge in their forum contributions. This aspect of moderation also relates back to other aspects: In relation to ‘information architecture’, moderators show other users where to find more resources. In relation to the ‘translation’ of information, moderators might have to translate expert knowledge for less knowledgeable users and know how to address different stakeholders such as players, scientists, developers, students, or educators. The Foldit moderator has to answer questions on a range of technical issues and is pointing users to additional resources or suggesting further steps. In what even the user asking describes as *„probably one of the stranger requests”*, the moderator even finds out that the background sound of the game is accessible for users who downloaded the client and describes how to find it: *„Great news! All the sounds are actually in the resources folder distributed with the game. They would be in the [Foldit Directory]/cmp-resources-[/resources/sounds . If anyone feels like making a cool remix out of them, that would be fun to hear too.”* In this case, the moderator might not have known the answer immediately but eventually did find out the answer to such an unusual question.

In Galaxy Zoo, moderators deal with many different questions about the objects classified. It is often up to them to answer a vague question about an image with appropriate information. This information can concern the project’s classification software (*„[...]Btw, almost every object in this image is a foreground star. Sometimes the software misbehaves and gets it completely wrong.”*), information about image processing (*„Here’s the edge where two images have been stitched together. The one to the left had some serious imaging problems”*) or astronomical topics, ranging from simple statements to long explanations of phenomena. Other aspects not directly related to the tasks of the projects also come up in questions and answers on the forum, including for example academic publications, related citizen science projects, or more general scientific discoveries not directly related to the project. The moderators have to deal with this vast range of topics and seem to collect a significant amount of knowledge about many different aspects of the citizen science projects. If they do not know an answer to a question themselves, they need to know where to point participants for further information, and whom to ask or inform if additional information or actions are needed.

## Part 2: ARTICLE: Exploring moderator roles in citizen science discussion forums

Citizen science projects engage citizens in scientific research, often involving volunteers in the collection or analysis of data (Rotman et al., 2012). Two big virtual projects, Foldit and Galaxy Zoo, allow participants to engage online by classifying images of galaxies (Galaxy Zoo) or by manipulating virtual representations of proteins in a competitive game (Foldit). Many different aspects of citizen science projects in general and these two projects specifically have been explored. However, there seems to be little research on the use of discussion forums in big virtual science projects. The role of moderators has been studied in other contexts, such as forums used in education or public forums about a variety of topics but seems to be overlooked in citizen science projects. Exploring how moderators act in discussion forums related to specific citizen science projects might offer insights on the overall communities of the projects and indicate how moderators can and do shape the interactions.

In this study, a selection of posts by moderators in the discussion forums of Galaxy Zoo and Foldit has been analyzed for a range of moderator actions. Some of the most frequent actions visible in the dataset can be categorized as social support, e.g. acknowledging or praising a user's post. Many examples of social support were found in both forums. In Galaxy Zoo Talk, pedagogical support (answering science-related questions and encouraging learning) was found in many of the moderators' posts. In the Foldit forum, technical support (helping users to run the game on their computers) was more prominent. The findings are discussed using Wiggins and Crowston's (2010) conceptual model of virtual organizations for citizen science. It seems that moderators take on a number of important roles in both Galaxy Zoo and Foldit. They can provide means of communication and connection between different stakeholders and they organize the contents of the discussion forums, including links to a variety of resources. Furthermore, moderators can offer participants answers to questions and support their learning; they seem to frequently acknowledge contributions to the project or the discussion forums, helping to increase participants' motivation.

# Background

Studies about citizen science projects discuss a variety of aspects, frequently including observations on possible learning outcomes for participants (e.g. Masters et al., 2016). Data from citizen science discussion forums has been studied for different purposes, most notably for examining user motivation and participation (Jackson et al., 2015; Raddick et al., 2010; Rotman et al., 2012) as well as the overall community dynamics (Luczak-Rösch et al., 2014). However, it seems that the different roles of participants and especially roles of moderators have not been discussed widely in the context of citizen science. Other literature on the moderation of discussion forums includes mostly data from forums used in an educational context (e.g. Kienle & Ritterskamp, 2007) or from publicly accessible forums with participants discussing general news, politics, hobbies or other common interests (e.g. Ghosh, Kale, & McAfee, 2011). It seems that studies both from educational contexts and from more public forums can offer insights about forum moderation that might also be applicable for discussion forums in citizen science projects. The following section gives an overview of the few remarks found in the literature on discussion forums or moderation in relation to citizen science.

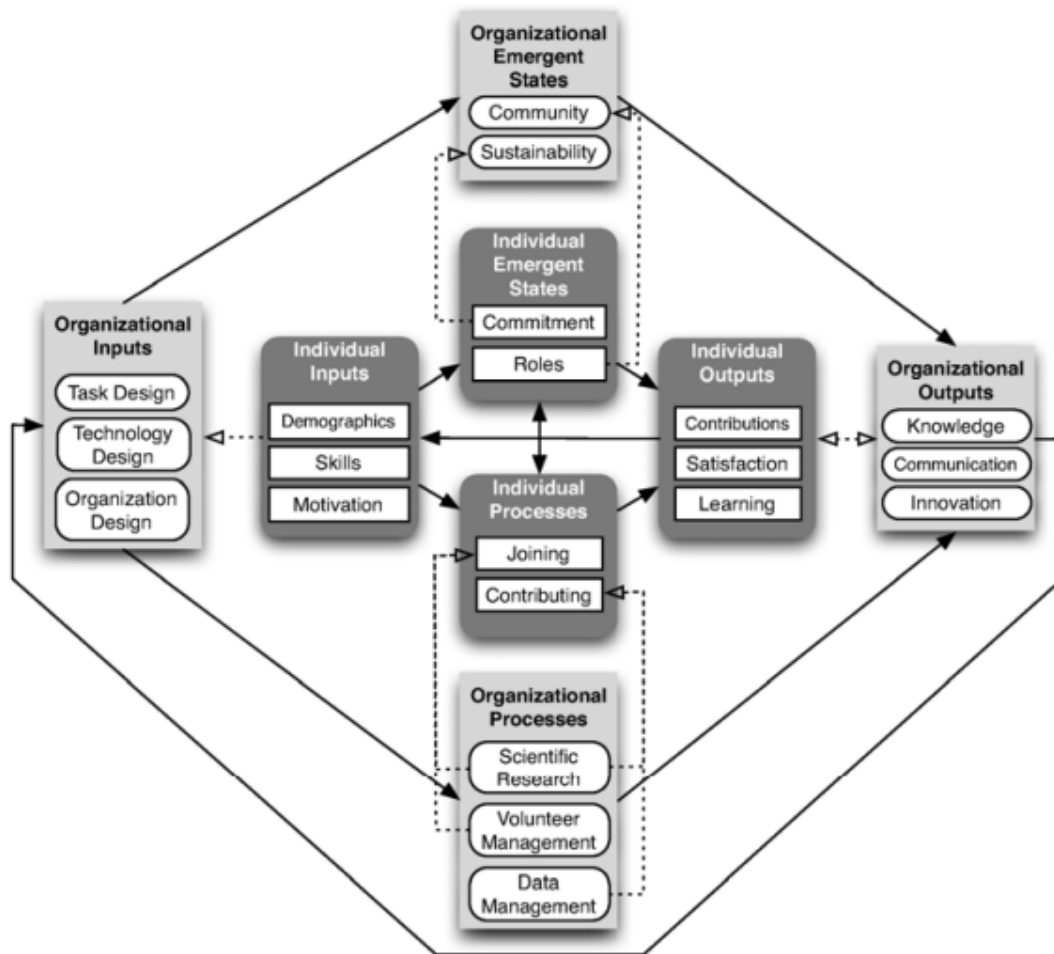
## Discussion forums and moderation in citizen science

Franzoni and Sauermann (2014) note that for collaborative citizen science projects, organizational practices are needed to facilitate problem solving and to structure information. They observe that „project leaders started to take on the role of moderators” (p. 7), synthesizing results and organizing threads in the ‘Polymath’ project. The role of the discussion forums in a Zooniverse project (Planet Hunters Talk) has been observed in a case study (Jackson et al., 2015). The authors argue that the functions incorporated in the Talk interface include responding to newcomers and supporting a feeling of community. Furthermore, they observe that „experienced volunteers act as moderators providing feedback and posting links for other volunteers” supporting „learning or curiosity” (p. 9). The study concludes that engaging in discussion in the forums and through personal messages can help volunteers to build identity within the project and might lead to more sustained participation. Similarly, Raddick, Bracey and Gay (2010) note that „some volunteers go beyond the basic task of classifying galaxies to other thoughtful interactions” like participation in forum discussions which might increase their understanding of science. All in all, only few studies mention discussion forums in citizen science, and notes on moderation in this particular context tend to be only short remarks included with other observations. Nevertheless, the few mentions indicate that the discussion forums might play an important role for the volunteers’ participation and feeling of identity in the projects. They can offer opportunities for communication and exchange, especially in projects that rely only on virtual collaboration or participation.

## Theoretical Framework

Since both technology-enabled citizen science and online discussion forums are relatively young phenomena, there are no established or widely used theoretical frameworks used for studying either. While there is research on different types of online communities, the theories used to study them vary with different researchers’ backgrounds and objectives. Additionally, the data collected from the two projects in this study, Galaxy Zoo and Foldit, can be considered from many perspectives. However, while there are different approaches of describing online communities, citizen science projects have some unique characteristics. A conceptual model of virtual organizations that takes these unique

aspects into consideration has been developed by Wiggins and Crowston (2010). Their model can help to explain the roles of individual actors and actions in a citizen science project and their relations to each other.



**Figure 4: A conceptual model of citizen science virtual organizations (Wiggins & Crowston, 2010)**

Wiggins and Crowston develop their model from previous work on crowdsourced and collaborative projects such as open source software development. They conceptualize the projects as work teams, arguing that a common output, rather than common practice is what characterizes the groups. Although it seems counter-intuitive at first to use models developed for small groups on massively large-scale projects, it can be a valuable way of describing the projects that tend to have a small core of involved members and a much larger user base that fluctuates a lot. The authors construct their model as an input-mediator-output-input model. This means that aspects of the project can be divided into inputs (e.g. the participants' skills), mediators (e.g. volunteer management) and outputs (e.g. participants' learning). Additionally, the outputs can serve as input for future mediators again (e.g. increased skills of participants that then join more or other tasks). The authors differentiate between two types of mediators, processes such as volunteer management and emergent states such as commitment. Furthermore, all four elements of the model (input, emergent states, processes, and output) can occur on both individual and organizational levels. The authors name examples for the different elements and indicate important relations by dotted lines (see figure 4). This includes for

example the commitment of individual contributors (individual emergent state) that might have a big influence on the overall sustainability of the project (organizational emergent state). Together with processes and other emergent states, and based on the inputs of the project, they will contribute to the outcome of the project both on an individual level (e.g. individual learning) and on an organizational level (e.g. knowledge that can be used for further research). In this study, the framework will be used to contextualize and discuss the findings on moderation from the two citizen science projects selected for this study.

## Methods

For this study, data from the online forums of two citizen science projects was gathered and analyzed. The following section gives some background information on the two projects, Galaxy Zoo and Foldit, and their respective discussion forums. It is followed by a detailed description of how the data was collected, processed and analyzed.

### Overview of the two citizen science projects

The two citizen science projects considered for this study are Galaxy Zoo, a citizen science project from the field of astronomy where volunteers classify images of galaxies, and Foldit, a citizen science project set up as a game where players manipulate protein structures. In their typology of different citizen science projects, Wiggins and Crowston (2011) name both Foldit and Galaxy Zoo as examples of ‘virtual’ citizen science projects. These are characterized by a top-down organization of activities, heavily relying on specially designed software systems, and conducting all activities in a virtual space without any additional physical elements. This means that virtual spaces for communication like online discussion forums are an important part of building up the community. Nevertheless, the two projects also have their own unique characteristics and present some interesting contrasts in relation to the use of discussion forums.

#### Galaxy Zoo

According to the project website, the first Galaxy Zoo project started in 2007 and asked volunteers to classify galaxies from images in some simple categories. The project has evolved since and has seen several relaunches and added different image sources as well as more complex categorization tasks. However, the main task for volunteers remains essentially the same: without any astronomical knowledge needed, users are asked to classify images by what is visible in them. (Galaxy Zoo, n.d.) This micro task relies on a large amount of classifications and having different members classify the same images several times to establish validity. The online discussion forums are one of the main communication channels among participants of the project. Additionally, a blog informs about recent developments. The forum is closely tied to the task of classifying images; images can be added to forum posts and discussions can be tied to specific objects (images). The discussions in the forum evolve quite slowly but there is a large amount of threads and posts by active users.

#### Foldit

Foldit emerged from the Rosetta@Home project where participants could use screensavers to contribute computing capacity to the task of protein folding. After users had voiced a wish to interact with the software, Foldit went online in 2008 and enabled users to manipulate the structures of protein representations in a game format (Franzoni & Sauermaun, 2014). This puzzle-like task can be seen as a macro task and has a very steep learning curve and only few participants become very talented top players. A special software client has to be downloaded in order to play the game. The Foldit online forum serves to help participants interact and contains many posts about troubleshooting the software. Only few users seem to make use of the forum; the discussions are moving slowly and the forum contains only a small number of posts and threads. However, it should be noted that Foldit also offers other ways of communication. Like in Galaxy Zoo, a blog informs about recent news from the Foldit project. Furthermore, a dedicated „Feedback” section on the forum deals with technical issues and

additional tools like an IRC chat can connect the players. Furthermore, some forum posts indicated that users even use external tools like skype or google hangouts to interact.

To conclude, the two citizen science projects considered in this study both use virtual communication systems (including discussion forums) to connect participants. However, the tasks and the communities differ from each other. The similarities and differences between the two forums will be considered further in the findings and discussion sections.

## Data collection and sampling

In order to explore the process of forum moderation in Galaxy Zoo and Foldit, data from the discussion forums was collected for analysis. Due to the different volumes of entries in the forums, the data was collected and selected slightly differently for each project.

An overview of the Galaxy Zoo discussion forum can be found on the project website<sup>1</sup>. Since the forum is large and collecting the data again would require a substantial amount of time, an existing corpus of data was used for the analysis. The data was collected in Mai 2015 using the Chrome browser extension „Webscraper”. Only data from the „Talk” discussions was used for the analysis, data from the older discussion forum was not included. Still, the dataset contained more than 36000 individual posts and more than 6000 posts by moderators, dated between September 2012 and Mai 2015. Therefore the dataset was narrowed down to the subforum „Science/Science”<sup>2</sup>. This left 857 posts in 138 threads. Out of these, 218 posts had been made by moderators (roles like „scientist” or „moderator” are indicated together with the user name in the forum). However, a first glance at the data revealed that 114 of the posts were in „sticky” threads supposed to give an overview and index of the forum. These posts were considered separately, leaving 104 moderator posts that were analyzed in detail.

The Foldit discussion forum is smaller and less active than Galaxy Zoo Talk, making it feasible to collect a newer dataset of the discussions. Only the main forum<sup>3</sup> was scraped for its content, while other sections (e.g. Archive, Feedback, discussions connected to specific groups or puzzles) were left out. Using a combination of the „Webscraper” tool<sup>4</sup> and an extractor built with import.io<sup>5</sup>, a total of 3980 posts in 768 threads was collected. The entries range from November 2007 to March 2016. The extractor tool had some difficulties with the nested answers in the forum and misread a small number of threads. The affected threads were identified by comparing the collected number of replies to the number of replies as stated on the forum overview pages, and then manually checked and corrected. On the Foldit discussion forum, staff members are marked with an additional icon (that did not show in the data, unfortunately). However, it is not visible who is performing which roles (e.g. moderator, developer, scientist) on the forum. Therefore, the analysis was limited to one currently very active moderator whose posts and behaviour clearly indicate responsibility for checking and guiding the discussions. This user has been active since Mai 2014; therefore the entire dataset was limited to posts made from this time onwards. „Sticky” threads were considered separately as in Galaxy Zoo (even though this concerns only few posts by the moderator in question). Finally, 91 individual posts were

---

<sup>1</sup> <http://talk.galaxyzoo.org/#/boards>

<sup>2</sup> <http://talk.galaxyzoo.org/#/boards/BGZ0000001>

<sup>3</sup> <http://fold.it/portal/forum/5>

<sup>4</sup> <http://webscraper.io/>

<sup>5</sup> <http://www.import.io/>



analyzed in detail for the Foldit discussion forum. Table 2 shows an overview of the data collected and selected for analysis.

	<b>Foldit</b>	<b>Galaxy Zoo</b>
Total dataset	3980 posts in 768 threads November 2007 to March 2016	36683 posts in 6031 threads September 2012 to Mai 2015
Limitation	Limited by time (from Mai 2014 on) 594 posts in 143 threads	Limited by subforum (Science/Science) 857 posts in 138 threads
Removing „sticky” threads	537 posts in 141 threads, thereof 91 posts by one moderator	701 posts in 133 threads, thereof 104 posts by three moderators

**Table 2: Overview of the data collected from the forums**

## Data analysis

The moderator posts in the selected datasets were first analyzed without a reference to frameworks, looking for common or recurring types of posts and possible categorizations. While some typical posts quickly stood out, finding common categories for the other posts was challenging. Therefore the data was analyzed using two different categorizations, one from an educational background that resulted from content analysis and one from an interview study on the moderation of technical help forums. While the first, educational study concerns the individual moderators’ actions on a very detailed level, the second interview study can contribute to get a more general impression of the overall set-up of the online forums.

Asterhan (2011) analyzed moderation styles in a synchronous discussion environment in a university course. Although the co-located setting and the synchronous discussion differ from the asynchronous online discussions in Galaxy Zoo and Foldit, the categorization by Asterhan can be useful in the latter setting, too. Asterhan distinguishes five moderator action categories:

- Pedagogical scaffolding support (encouraging learning)
- Interaction support (encouraging participation)
- Managerial support (task design, task completion, monitoring)
- Moving forward (offering discussants a new perspective)
- Involved participation (joining the discussion as an equal participant)

Furthermore, Asterhan mentions two more categories that did not occur in the data from the co-located setting but might be found in other settings:

- Social support
- Technical support

These seven categories can be used to describe the individual actions that moderators take in their different posts. The categorization is combined with additional data in Asterhan’s work to arrive at a set of different moderation styles that can be distinguished by the content and meta-data of the posts and threads. However, since the dataset used in Asterhan’s work included more information (moderators had to select a type of answer before posting), the analysis of the two forums in this study will differ from Asterhan’s approach.

In addition to Asterhan's framework that originated from an educational setting, observations by Frith (2014) are also considered to describe the overall set-up of the two online forums. Frith's work is based on technical help forums and distinguishes five roles that moderators can take on:

- Moderators as knowledgeable non-technical experts (moderators do not have to be experts in the topic of the forum, but do need some special knowledge for their actions)
- Moderators as quality control experts (controlling spam)
- Moderators as translators (translating between different levels of expertise)
- Moderators as information architects (organizing the overall structure of the forum, e.g. subforums, „sticky” threads)
- Moderators as tone-setters (determining the overall tone of the forum, including how welcome newcomers are to the forum)

The datasets of 91 moderator posts (Foldit) and 104 moderator posts (Galaxy Zoo) were coded using both Asterhan's and Frith's work. As Asterhan noted, social support might be very subtle and intertwined with other posts, suggesting the need for a dual analytic system. Furthermore, neither categorization was explicitly created as a framework for analyzing moderation in the citizen science discussion forums. Therefore, the categories were not used as mutually exclusive categorizations. Instead, all possible applicable categorizations were indicated for each post.

# Findings

The study of moderation in Galaxy Zoo and Foldit found differences as well as similarities between the two forums. In the following, the context of the two forums and some numerical descriptions will be presented, before moving on to the detailed description of moderator actions and moderator roles in the forums.

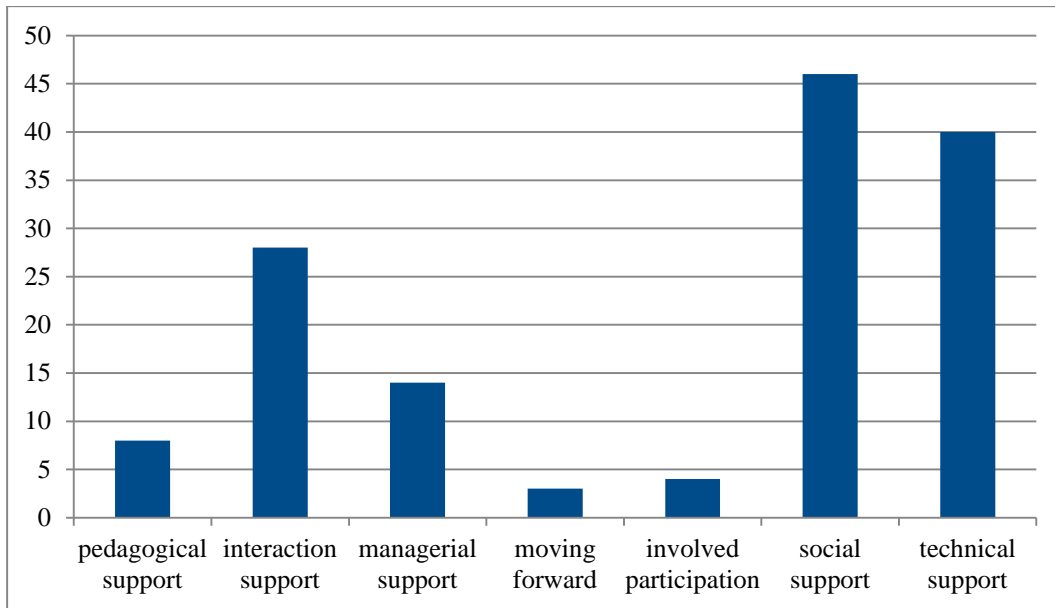
## Context of the two forums

Some general numerical descriptions of the forums can give an overview of the two projects. A full overview of the numerical data collected from the original datasets and the selected datasets for analysis can be found in Appendix 1. In both forums, threads tend to be short (70-80% containing less than 6 posts) with only few exceptionally long threads (more than 20 posts, highest percentage in the subforum of Galaxy Zoo with 6,5 percent). Also, a large percentage of users (72-78%) only contribute once or twice to the discussions. It seems that Galaxy Zoo has a smaller percentage of core users (about 15% of users contribute to 80% of posts) than Foldit (about 30% of users contribute to 80% of the posts). The percentage of users with more than 10 posts is the same in the Foldit original dataset and the selection (7%). However, in Galaxy Zoo, the subforum has a higher percentage of frequent posters (14%) than the forum overall (5%), suggesting that the subforum might have more active users. Additionally, posts in the subforum also tend to be longer than in other parts of the forum. Ranking the users by the number of their contributions and considering the top 25 users shows that among the most frequent users are moderators (1 in Foldit, 3 in Galaxy Zoo), other staff members (3 in Foldit, 4 in Galaxy Zoo), and highly involved users (6 of the 25 most frequent users on Foldit are ranked among the top 100 players by the games' scoring mechanism).

To get a first overview on the moderator involvement in the two forums, data about different aspects of the forum was collected and compared. Between 40 and 55% of the threads have moderator contributions. While moderators rank as the most active users, they make up between 17 and 25% of the forum posts. This indicates that the moderators are the most active users over time, but different users together make up a larger part of the contributions. It is noteworthy that moderators tend to start new threads much less than the average of other users (9 of 134 threads started by moderators in Foldit, 8 out of 138 in Galaxy Zoo). Many of the threads started by moderators are so called „sticky” threads (threads that stay visible at the top of the forum at all time, regardless of the otherwise chronological order). In these threads, moderators (and other staff members) are very active. In the Galaxy Zoo subforum, one moderator in particular stands out by contributing almost all of the posts in sticky threads. The same moderator also shows much higher activity in this subforum and on average much longer posts than the other moderators, indicating a special interest for this particular subforum.

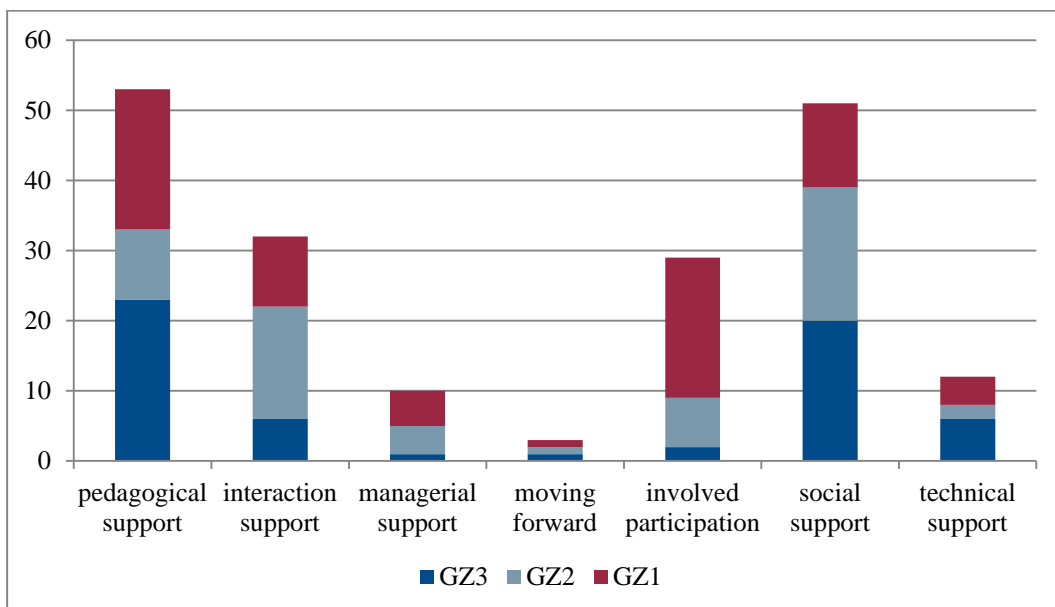
## Moderator actions

The actions of the moderators on both forums have been coded for the selected individual posts using Asterhan's (2011) framework of seven types of moderator actions. Figures 5 and 6 show overviews of the different moderator actions found in the two discussion forums. It should be noted that the categories have not been used exclusively, therefore they do not add up to the total number of posts.



**Figure 5: Foldit moderator actions (total of 91 posts analyzed)**

In Foldit, technical support and social support are very dominant. For Galaxy Zoo, the actions are also broken down by the individual moderators, here named GZ1, GZ2 and GZ3 (in Foldit, only one moderator was observed). In the selected dataset from Galaxy Zoo, the most dominant categories were pedagogical and social support. The three moderators in this dataset contribute differently to the categories, indicating some preferences or specializations. In the following, all of the categories are explained in detail and with additional examples. All examples below correspond to complete posts, with omissions indicated by square brackets; links to other resources (external or internal) have been replaced for better readability.



**Figure 6: Galaxy Zoo moderator actions (total of 104 posts analyzed)**

## Pedagogical support

Generally, pedagogical support on both forums consists of either a full explanation or answer to a specific question, or pointing users (generally or for specific questions) to additional resources. Additional resources can be found both within the project (e.g. other forum posts, blog posts, wiki [Foldit], posts from an older forum [Galaxy Zoo]) or outside the project (e.g. academic publications, videos, field-specific websites). There is a notable difference in pedagogical support (i.e. encouraging users' learning) between the two forums. In the Foldit discussions, the moderator posts often thank other users for contributing links to relevant information, but rarely are actively promoting any additional resources. One exception is a post about an online course but even in this post, the moderator points out that it had been found by another user: „*Susume found this great online course! [...] that might be of interest to Foldit science fans. [...]*”. Additionally, a few moderator posts answer questions of educators trying to set up Foldit for a class of students. Although the nature of these posts is more technical than pedagogical, they are still noteworthy, showing a connection to a different group of stakeholders concerned with education. In discussions from Galaxy Zoo, on the other hand, moderators seem much more active about answering questions or pointing users to additional resources. Examples range from very simple answers („*If you mean the blue blobs, these are star forming regions in the galaxies*”, GZ3) to long explanations with many links to scientific resources (mostly posted by moderator GZ1). Explanations deal with topics of astronomy (e.g. certain types of galaxies or stars) as well as technical issues of the images (e.g. optical artifacts resulting from image processing). In one post, moderator GZ1 acknowledges that the different explanations and resources are not equally easy to understand: „*There is a quantum leap between easy and hard stuff, and not much in between [...]*”.

## Interaction support

Interaction support (i.e. attempts of directing the social interaction, encouraging participation) is not standing out as a main moderator action in either forum. Nevertheless, the moderators do direct the interactions to some extent. For Foldit, this often occurs in connection with technical support: members are asked to give more information on the problem in the feedback section of the forum: „*This is a pretty valuable bit of feedback I'd encourage you to cross post in the feedback section along with your screenshot. If others are having similar issues and also chime in with their screenshots, then we are able to investigate the issue better.*” While participants are often addressed with their user names in posts that thank for or answer to specific posts, requests about the interaction are often much more general or include a more broad request after an initial specific answer. In some cases, the moderators do address specific users and request more information: „*I hadn't seen that, do you have a link for us?*” (Foldit), „*An update? I think a few of us have been checking in to see how things are going... ;-)*” (GZ2). Other actions that can be called interaction support include issues of organizing the forum threads. An example from Foldit is: „*Valid suggestions but we're veering a little off here from puzzle balance to recipes. :) Let's get these perfectly great ideas into their own thread or a feedback suggestion, and not lost in the shuffle of the overall picture. Thanks!*” On Galaxy Zoo, a user noted that one of the moderators knows Dutch and started writing in Dutch, receiving the following answer: „*Actually, we try to keep things in English, so everybody can understand it. [...]*”. All in all, directing the interactions on the forum is not a highly visible task on the forums but examples of it can be found, asking users in general or specific users for (additional) information, or managing forum threads.

## Managerial support

Asterhan (2011) describes managerial design as task design, task completion and task monitoring. In her research, the forum discussion was the actual task, whereas the online forums for Galaxy Zoo and Foldit only serve as support for the main tasks of classification or gameplay. Therefore, examples of managerial support that actively support the ‘task’ of playing or classifying are difficult to find in the forums. To some extent, announcements about the project can be categorized as managerial support, e.g. *„You’ll be happy to know a new Ebola puzzle is on the way this week. [...] We have run a few Ebola puzzles in the past, and we know that it is currently a matter of great interest to our community.”* (Foldit). Additionally, the overall management of the forum could be seen as managerial support, too. In this case, there would be some overlap with the previously mentioned examples of interaction support but also some issues that don’t address any users but the organization of the forum threads *„I am more than happy to ‘sticky’ it to the sidebar one you have created the thread :-)”* (GZ2). Generally, the „sticky” threads mentioned in this post might be seen as managerial or pedagogical support; however, they were excluded from the detailed analysis.

## Moving forward

In her study, Asterhan (2011) identified „moving forward” as an additional category of moderator actions. However, in the case of Foldit and Galaxy Zoo, discussions are not aimed at leading towards a solution like in the educational setting, making it less necessary for moderators to move the discussion forward or intervene with „pivotal turns”. In Foldit, the moderator in one case offers an own opinion on the idea of rewarding players: *„Instead of paying people, why not band together and do something charitable instead, such as forming a team for a quality organization like Extra Life?”*, which could be seen as moving forward (or as involved participation). There were not many controversial threads observed in the selected datasets, which might also lessen the necessity for moderators to intervene more actively to move a ‘stuck’ discussion forward.

## Involved participation

As for involved participation, the lines between this category and others are very blurry. Other users on the forum also answer scientific and technical questions, and acknowledge other users’ contributions. It is difficult to tell which actions are specific to the moderators only, even if they might be frequently be done by moderators. There are some examples of moderators sharing personal information or their own opinions; however, the moderators do not seem to be as involved in the tasks as other users. The moderator GZ1 is an exception to this, frequently posting own observations and also questions on the forum: *„Thank you Dr Simmons. I had noticed that some of these compact sources had little blue cores, and I had wondered if that was significant.”* Other actions could also be seen as social support or pedagogical support, e.g. asking for a clarification also for other users’ sake or acknowledging users’ actions: *„Wow! Exciting stuff. First question - why is it so important to observe / find blue ellipticals?”*

## Social support

Examples of social support are very prominent on both forums. This includes acknowledging users’ contributions and thanking them for information or praising interesting finds. Some examples of personal messages can be found in the forum, additionally, personal issues might also be conveyed through other channels. On Foldit, many posts categorized as social support thank users for sharing a link in the forum; they are often very short and mention the username of the contributor (*„Thanks jeff101!”*) but can also be slightly longer: *„Great link! Thank you for sharing it with us. Looking*

*forward to reading more cool news in this area.*” Similar messages also occur on Galaxy Zoo: *„Nice video. Tx for posting ! :D”* Additionally, publications based on data from the project are also commented on frequently (e.g. *„Excellent work, Brooke and company! :-D”*). In one post, a moderator even points to a list of all the contributors: *„The original paper - congratulations to all the discoverers (see pages 5-6) - Thanks to @Stellar190 for posting the link on social media. For a list of all the contributors to the voorwerpje mini project looks at pg 22 :-) [...]”* Another post about a scientific publication is edited with the following information: *„[...] More importantly here is a full list of all the contributors! Congrats everyone :-)”* These examples suggest that the forum moderators play an important role in social support not just for the forum but for the overall project, acknowledging and celebrating successes, and helping volunteers to see what happened with their contributions.

As Asterhan (2011) notes, social support can be very subtle. This can also be observed in these two forums: moderators often start their messages with a greeting, mention user names in answers, use emoticons, or phrase requests indirectly. Furthermore, social support can be intertwined with other messages, for example in answering a newcomer’s question in Galaxy Zoo: *„Hi KristinnFrans and welcome to the Zoo Are you talking about this object ? : AGZ0004cf0 That is a star from our galaxy, but the colors are not real. More here : [link] Happy hunting ! :D”* Another time, a moderator dampens a (perhaps disappointing) answer by adding that the image is nevertheless intriguing: *„I think this is just one star. Pretty impressive though !”*

## **Technical support**

Examples of technical support are very prominent in the dataset from Foldit. They are less common in the dataset from Galaxy Zoo, however, this is likely due to the selection of the subforum; other subforums might include more technical issues. As mentioned previously, the technical support Foldit is often connected to interaction support: asking users to give more information about the issue: *„So sorry to hear you’re having trouble, if you could provide a bit more information, or perhaps a DXDIAG, that may help us narrow down your issue. Thanks, we want to get you up and folding again in no time flat!”* The moderator seems to adapt the language and advice for different users. This also includes educators who want to use Foldit with an entire class; they often receive extensive information and feedback from the moderator. While the issues on Foldit are often very specific and allow moderators to only point to possible solutions or more needed information, the few examples of technical support on the Galaxy Zoo subforum deal with more general issues: *„You will be able to participate with your Zooniverse account. It is a Zooniverse project after all ! :-D”* (giving information for a sister project) or *„The UKIDDS images were completed last Tuesday - it caused the site to temporarily go down. It was Michael Parrish who made a post in the bug thread.”* (mentioning technical difficulties in an explanation). In some cases, the moderators can only acknowledge technical difficulties and not offer a solution or further steps: *„We’re looking into it here further and appreciate you bringing it to our attention!”* (Foldit). Those examples could also be considered social support. Although they deal with technical issues, they do not actually solve the problems or offer users any further steps but acknowledge the problem and state that further actions will be taken eventually.

## **Moderator roles**

The data also shows examples of posts that match Frith’s (2014) identified moderator roles resulting from interviews with moderator roles. The only exception is role of „quality control experts” that is not visible in the data – which is not surprising, since the removal of spam will not be visible in the

forum posts collected. The „sticky” threads, especially by moderator GZ1, are great examples of what Frith describes as „information architecture”. The moderators bring together information from different sources, creating an overview. Frith’s concept of „moderators as tone-setters” is closely related to Asterhan’s (2011) moderator action of social support. However, the idea of „moderators as translators” differs from Asterhan’s categorizations. No examples of ‘translations’ between different users were found in the datasets. However, some form of translation or mediation can sometimes be seen: Moderators serve as a connecting between the volunteers and the researchers or developers, communicating information between the different groups. Finally, Frith observed that moderators are not necessarily experts in the forum topic, but nevertheless very skilled and knowledgeable („knowledgeable non-technical experts”). In both datasets, moderators demonstrate a lot of knowledge about various aspects of the citizen science projects, even though they are not necessarily trained experts. In particular, they know where to find additional help or resources and frequently point these out to other users.



# Discussion

## Comparing the two forums

Since the criteria for selecting a subset of data was different for the two forums, comparing the data should be approached with caution. The lack of technical support found in Galaxy Zoo Talk for example likely results from the selection of the subforum, rather than a difference between the two forums. However, some general observations about the similarities and differences of moderators' interactions with other users in the two forums can be made.

Considering the general forum setup and activity, it seems that Galaxy Zoo Talk is much more tied in with the Galaxy Zoo citizen science project than Foldit with its respective forum. There are many users and contributions in the discussions and the set-up of the forum is closely connected to the classification task. The forum seems to be one of the main communication channels of the project. Foldit, on the other hand, offers the forum as one out of many communication channels. Additional data might be found in the chats or in forum posts directly tied to individual puzzles or groups (these were not part of the analysis), but the content immediately accessible from the public forum seems less tied in with the task, that is performed in a separate software setup. This might explain the quite low activity on the Foldit forum, in addition to a smaller project scope and the goal of retaining mainly talented and specialized top players, whereas Galaxy Zoo aims to maintain a broad user base.

In both forums, social support was very prominent and could be found both in individual posts and more subtly in the tone and character of other posts. However, it seems that the moderators on Galaxy Zoo were more active in providing pedagogical support, supplying users with answers and explanations on various levels of difficulty, as well as links to internal and external resources. In Foldit, such resources seem to be posted mostly by members and then get acknowledged by moderators in form of social support. The posts categorized as social support on Galaxy Zoo show a mostly praise and acknowledgement of individual users' contributions and classifications. The comments range from acknowledging aesthetic qualities of images and funny or interesting looking images, to celebrating users' contributions to scientific publications. The higher level of acknowledgement of the task for Galaxy Zoo might be related to the projects' set-ups: Foldit is a very competitive task and users can gain recognition in form of the game rankings visible on the website. For Galaxy Zoo, the forum might offer a form of recognition that is not offered by rankings or achievements like it is possible in the Foldit game. In both projects, there seems to be a lot of interest for the publications that result from the projects. More examples of this were found on Galaxy Zoo than on Foldit, however, this might be due to the different forum sizes and the selection of the subforum particularly concerned with science.

## Moderators within virtual organizations of citizen science

### Moderators as individuals

Within the conceptual model of virtual organizations in citizen science by Wiggins and Crowston (2010), moderators can be seen as individuals who bring their own input (e.g. skills, motivation, demographics) into the project. Their commitment and the roles they take on (emergent states) as well as their contributions (process) will influence the discussion forum and possibly the entire project, impacting the organizational emergent states (community, sustainability) and processes (especially

volunteer management) as well as individual outcomes (e.g. learning). With their special role, moderators likely have a particular influence on the overall organizational outputs. Examples show that moderators might collect and organize knowledge from a variety of sources for their own and other's benefits. These collections can be seen as an individual outcomes (collected by individual moderators) or as organizational outcomes (making the collected information publicly available). Furthermore, moderators have a big influence on the communication in the community, since the discussion forums are one of the most visible and most accessible communication channels.

### **Connecting the individual and the organizational level**

In addition to seeing moderators as individual participants of the project, they can also be seen in the function of connecting different stakeholders. Moderators often seem to 'translate' or mediate between the volunteers on the one hand and the scientists and developers on the other hand. They possess an overview of the system and can relate information to the different parties concerned. The data from the forums show examples of collecting and presenting information from individuals to the development team. On the other hand, moderators also convey information from the organization to the volunteers, including information on new or changed inputs (such as changes in technology design) as well as project outputs like overall statistics or publications. Research on a different project found that even in projects with little opportunity for participation beyond a simple task, interactions between researchers and citizens might still be possible through online discussions (Savio, Prainsack, & Buyx, 2016), which can underline the importance of moderators' activities in these discussion forums.

To some extent, moderators on the discussion forums seem to belong to the two levels of the model at the same time: in some cases they participate much like any other volunteers and refer to 'the team' or 'the developers', in other cases they speak for the entire project as 'we'. Although the discussion forums are primarily a communication channel and do not represent the entire project, the moderators do influence the part of the community that is visible on the forums (which might also include participants who only read the discussions). The moderators set the general tone of the forum and build structures for the information collected. They provide a vast range of support, including technical support and social support, ensuring that the project runs smoothly.

### **Finding, collecting and presenting knowledge**

In both forums, the moderators show a vast range of knowledge about different topics related to the task, even though they are not necessarily trained experts. They seem to be especially good at pointing other users to the right resources, and contacting the right persons for further questions or suggestions. Roth and Lee (2004) suggest that scientific literacy can be seen as the ability to „be competent in finding whatever one needs to know at the moment one needs to know it” (p.8). The moderators themselves show great skill of finding relevant information. Furthermore, they also acknowledge when other users contribute additional resources or information, recognizing and encouraging this particular notion of literacy.

In Wiggins and Crowston's (2010) model of citizen science projects, knowledge is considered as an organizational output. It is collected and published in the project databases or academic publications. However, increased knowledge can also be gained by the individual participants. Their knowledge is in some cases also collected and published in blogs, wikis, or discussion forums. Some moderators present overviews in „sticky” threads in the discussion forums. This collection of knowledge from the

discussions and from other resources becomes more accessible and more visible, moving the individual contributions to a more general, organizational collection of knowledge.

## Increasing satisfaction and learning

Two elements that Wiggins and Crowston (2010) describe as individual outputs are prominently affected by moderation: Data from the Foldit discussion forum and Galaxy Zoo talk suggests that moderators can have an influence on the satisfaction and the learning experience of participants.

Firstly, the prevalence of social support might increase the participants' satisfaction or motivation. The emphasis on recognizing their contributions, both to the project and to the discussion forum, indicates that moderators aim to acknowledge these contributions, possibly increasing participants' satisfaction with the project. Rotman et al. (2012) argue that external factors might influence participants' motivation over time and mention learning opportunities, the recognition of contributions, and the feedback about what happened with the data, as examples for these factors. The data from the Foldit and Galaxy Zoo discussion forums shows that moderators might have an important influence on these particular factors, making them influential for the users' satisfaction (output from participation) that can also serve as an input (in form of motivation) for joining more or other projects and continuing participation.

The second prominent element in the data on moderation is learning. While a lot of pedagogical support was found in the Galaxy Zoo Talk data, the Foldit moderators mostly seem to acknowledge the support that other users give. It seems that Galaxy Zoo users are more likely to be provided with direct answers by the moderators, while Foldit moderators encourage users to seek information by themselves (also offering other communication channels for that purpose). Wiggins and Crowston (2011) comment on the irony that especially the big, academically driven virtual citizen science projects tend to include „relatively little provision of educational materials for participants“. However, discussion forums might help to provide these materials for and also by users. Moderators can play a special role in curating the knowledge collected from users as well as other sources and in acknowledging or actively encouraging learning about the science behind the tasks. Moderators can also serve as translators or mediators between different project stakeholders, opening up opportunities for further discussions and learning. In Foldit, the game itself has a steep learning curve and forum discussions with moderator involvement seem to discuss mostly technical problems, but also scientific issues. In Galaxy Zoo, classification is a straightforward task but users nevertheless seek out the forums and are supported by moderators in appreciating images, learning about astronomical facts and learning about the technical details of the images processed.

One moderator from Galaxy Zoo points out another interesting aspect of learning, mentioning a „quantum leap“ between easy and hard explanations of science. This suggests that there is an interest for explanations on an intermediary level that are detailed but can be understood without too much scientific background knowledge. Citizen science projects might be a good starting place to fill this gap, either formally by providing materials, or informally, by building up resources like discussion forums or wikis. A study on motivation and participation in citizen science also suggest that „projects should provide authoritative resources as volunteers seek more knowledge, access to sustained members, and highlight how volunteers may take on additional roles once discovery is made.“ (Jackson et al., 2015, p. 9) Even if this information is not given by project officials, moderators might

be able to provide it. Additionally, their acts of social support might also encourage participants to learn more or to share their knowledge.

## Conclusion

Analyzing a selection of moderator posts from the discussion forums of two citizen science projects (Galaxy Zoo and Foldit) revealed a range of possible actions and roles that moderators might take on. Most notably, moderators in the analyzed discussion forums seem to be active connecting different project stakeholders and curating the participants' and other sources of knowledge. Beyond providing resources for learning, they also offer answers to users' questions and encourage them to learn more. Furthermore, they might have a big influence on the satisfaction of the users and their motivation, offering recognition for activities that are not the main task of the projects (i.e. providing learning resources for other users in Foldit or finding aesthetically pleasing images in Galaxy Zoo) and informing users how their contributions is used.

The increase of citizen science projects in many areas makes it important to consider different aspects of these complex set-ups involving many stakeholders. Analyzing how forum moderation is used can help to identify further opportunities for research and improvement of citizen science projects. It seems that moderators take on a range of activities that help the projects to maintain and support the community and to help participants learn about the subject. It might be interesting to consider how far their roles are wanted, recognized, or encouraged by the project developers.

Although the two projects have different tasks and different community set-ups, they are both categorized as virtual projects by Wiggins and Crowston (2011). Even between Foldit and Galaxy Zoo, there seems to be a difference in the range of communication channels available to and used by participants and in the focus of moderator activities. In projects that rely on physical elements more than information technology, discussion forums might be used differently or not at all. Nevertheless, it seems that the roles identified for the moderators in this study, i.e. connecting stakeholders, encouraging learning, organizing information, and recognizing users' contributions, are likely to be needed in other citizen science projects as well. Considering by whom and how these tasks are fulfilled might give valuable insights about citizen science projects and help to design projects that can sustain participants' motivation and offer them learning opportunities.

## **Part 3: DISCUSSION AND CONCLUSION OF THE ARTICLE**

The following sections contain a summary of the discussion from the above article and more detailed discussions about aspects of forum moderation in citizen science projects. These include a discussion about the relation of citizen science projects and their discussion forums to other contexts, observations on citizen science forums as an opportunity for learning, and aspects of forum moderation that seem to relate to learning. Additionally, some limitations of the study and challenges encountered in the research projects are presented, along with opportunities for further research that result from them.

## Discussion

As discussed in the article above, moderators can fulfill a number of important roles within citizen science projects. Although the focus of activities seems to differ between Galaxy Zoo and Foldit, some roles stand out as particularly relevant for citizen science. Firstly, moderators can help to connect the volunteers with the researchers or developers of the project. As participants or non-experts, they can help to translate and transfer information between different stakeholders. This includes especially the communication between project developers and researchers on the one hand, and participants or volunteers on the other hand. Furthermore, other stakeholders that might have additional or special concerns also might receive support from moderators. In Foldit, examples were found of educators requesting help with setting up the game for a class of students, revealing that educators and students might be additional stakeholders for the Foldit project that may not be immediately visible. Additionally, moderators navigate a vast range of knowledge and resources. They might make some of this knowledge visible and available to others by providing comments and collections, often in the form of „sticky” threads. This includes not only factual knowledge or answers to specific questions, but also strategies for finding answers, resources that can provide more information, and information about the level of difficulty of understanding these resources. Furthermore, moderators might influence the participants’ satisfaction with the project and their motivation to keep contributing by means of social support. Moderators frequently praise users’ contributions to the discussion forums or to the project in general and can offer participants feedback about the results from their contributions. Finally, moderators also seem to have an influence on learning, offering different forms of direct pedagogical support as well as other supporting activities that might also encourage or help participants who want to learn about the scientific backgrounds of the project.

In addition to the four moderator roles discussed in the article, the data collected from the discussion forums can also shine some light on a few more questions. This includes the comparison of citizen science projects and their discussion forums to educational and other settings, and more detailed discussions about the role that online forums in general and moderators specifically can play for encouraging and supporting citizen scientists to learn more about the research and topics they are contributing to.

### How does citizen science compare to educational and other settings?

The literature review on forum moderation revealed that there is a gap in the research between formal educational discussion forums on the one hand, and more openly accessible public forums for a range of different interests on the other hand. It was suggested that citizen science discussion forums might lie somewhere in between these two variations: They have implicit or explicit learning goals and a very formal and hierarchical core of project developers and scientists that determine the tasks for participants, likening them to discussion environments in educational settings. In selected posts from Galaxy Zoo and Foldit, it seems that there is very little controversy or control of inappropriate content, and that moderators instead focus on praising contributions and encouraging learning.

However, the volunteer members of the ‘community’ are not determined by formal setups and participation is open to anyone; many participants contribute only sparsely and over a short time (Franzoni & Sauermann, 2014), similar to more public forums or open source software development. Exploring the many ways of conceptualizing virtual communities revealed that a similar ‘in-between’ situation exists also for theoretical frameworks: while citizen science projects cannot be characterized

as learners' communities, describing them as communities of interest might not be entirely suitable either. A 'community of purpose' might be the best term for the overall project. The discussion forums of citizen science projects are only a small part of the project, members might not participate in them or projects might not even offer discussion forums. Nevertheless, it seems that even the entire community or virtual organization of a citizen science project falls in a context in between educational learners' communities and other discussion forums that can be characterized as communities of interest. Therefore, theories from either of these related contexts might be applicable for citizen science as well, if they are applied carefully and with consideration to the unique characteristics of citizen science projects.

Using categorizations from an educational forum as well as technical help forums showed that some categories found in those contexts are also very prominent in citizen science discussion forums, whereas others seem absent from the datasets. For example, it was difficult to find examples for Asterhan's (2011) description of managerial support, since the forums are not used for the main 'task' of the projects but for additional support. Nevertheless, many other categorizations both from Asterhan's work as well as Frith's (2014) descriptions of moderator roles could be found in the data easily, indicating that the niche between formal education setups and very loosely connected virtual communities might in fact be the right and appropriate way of contextualizing citizen science projects in general as well as their corresponding discussion forums.

## Can discussion forums support learning in citizen science projects?

Citizen science projects are often described as an opportunity for informal education and learning and many studies consider participants' learning experiences (e.g. Brossard et al., 2005; Crall et al., 2012; Jordan et al., 2011; Trumbull et al., 2000). There are differing findings about the extent of learning that participants might experience, depending on the participants' previous knowledge, the setup of the project and the method used for analyzing knowledge or attitude changes. However, many authors find that participants do indeed increase their knowledge, especially knowledge directly related to the projects' specific tasks. Although citizen science is often promoted as a valuable learning experience, especially the 'virtual' projects stand out with their lack of educational materials provided to help participants learn (Wiggins & Crowston, 2011). However, in these projects, participants might still learn about the science behind the tasks if they engage in additional activities, like discussion forums (Luczak-Rösch et al., 2014). The extent and characteristics of support for participants' learning differed largely even in the very small sample of data analyzed for this study. Especially if no official educational materials are provided, the activities of moderators and other users in the discussion forums might have an influence on the information accessible and visible to volunteers: In Galaxy Zoo, one moderator constructed large collections of helpful resources, and all three moderators provided answers or resources to specific information. In the Foldit discussion forum, the moderator acknowledges users who post information from other resources but is not actively contributing additional information. These differences are likely related to the forum setup and the moderators in question: the Foldit moderator is not a top player of the game and might not be very interested in additional scientific information related to protein folding. The Galaxy Zoo moderator who contributes many resources is clearly very involved in the project and is an active participant as well as moderator, interested in the astronomical and technical phenomena related to the classified images.

The entire discussion forums with contributions by moderators and other users can be a valuable resource for participants who are interested in learning more about the tasks and scientific



backgrounds of a citizen science project. However, only a small part of the participants seem to be active contributors to the forums and it is difficult to tell if they might have developed similar interests even without the moderators' support, without the discussion forums, or without the citizen science project in general. Nevertheless, the moderator posts analyzed in this study do show that online forums of citizen science projects to offer a point of contact between citizens and scientific research. Moderators are able to offer participants' information and support, providing a range of different answers and resources appropriate for different levels of understanding and previous knowledge. Furthermore, the social support given by moderators, acknowledging and praising users' contributions to the forums and to the overall project, might also help to motivate users to become or remain engaged in the projects, to learn more about the topic or science in general, and to share their knowledge with others.

## How do moderator roles in citizen science projects relate to education?

The moderators of the discussion forums studied seem to have different connections to learning and education. On the one hand, they themselves can be seen as learners or „knowledgeable non-technical experts”, navigating a vast range of information resources. On the other hand, moderators seem to have a particular impact on participants' learning by offering different kinds of support activities.

### **Moderators as learners**

The moderators have to navigate a vast range of knowledge and resources in order to successfully support participants. They might create overviews of information from different sources and need to know where to find answers to specific questions posted by users. Even though they are not necessarily trained experts in the scientific field, they nevertheless exhibit a lot of knowledge about the project tasks, related scientific concepts, and technical issues. They also exhibit curiosity and might act as participants by themselves, showing what Asterhan (2011) calls 'involved participation'. Furthermore, moderators seem to use their own curiosity in combination with other moderation activities. This includes for example stating that they provide a certain contribution for their own and other users' sakes, sharing resources they themselves found helpful or interesting, or asking questions that other users might also be curious about.

### **Moderators as educators**

The most obvious relation of moderators and participants' learning are instances of pedagogical support identified in the dataset. These include short answers and long explanations about different scientific and technical concepts and issues. Furthermore, moderators might point participants to useful additional information both within the project and in external resources. These activities can be seen in offering direct support to individual users and also in providing information targeted at a more general audience. The literature review showed that in bigger forums, shaping the overall architecture of the information might even call for automated or technically supported moderation to avoid bad content and highlight valuable contributions to the forum. In relation to this aspect, the selected posts showed more resemblance to educational forums with a focus on encouraging participation and supporting learning rather than trying to control the quality of contributions. However, even highlighting valuable contributions can be useful for organizing the discussion. In the Galaxy Zoo Talk and the Foldit forum there are no mechanisms for evaluating contributions, leaving it up to moderators (or committed volunteers) to collect and present overviews of useful information. The large amount of information on Galaxy Zoo Talk in combination with an older forum seems to require particular attention: one of the moderators creates long and detailed lists about where to find what kind

of information. Furthermore, moderators also explicitly or implicitly point how easy or difficult to understand some of the resources are. They seem to adapt their explanations for different user groups and provide not only the resources as such but also information about their accessibility to users. Additionally, social support might further encourage users' learning: moderators frequently acknowledge or praise users' contributions to the forums or to the citizen science projects. In addition to task-related information, moderators sometimes also offer insights into the process of scientific research, e.g. by informing users about publications that resulted from their data. This might serve as encouragement and also help users to learn more about science and research in general.

## Challenges, limitations, and further research opportunities

The background of the study, the categorization of the selected moderator posts, and the selection of the projects and the data all had a big impact on this thesis. These aspects raised challenges for the research and present limitations for the results. However, they might also open up new questions and opportunities for further research.

One of the most challenging tasks for this thesis was to find and decide upon relevant background information and appropriate theoretical frameworks and methodology. The study brings together two topics that both include a variety of different aspects and backgrounds: Research on forum moderation can be conducted from an educational perspective, it can consider aspects of communication and of technology, and it often includes the additional topic of interest that is discussed in the forum. Research on citizen science also often includes the particular topic or field of the project; additionally, it may consider the backgrounds, motivations, or learning outcomes of volunteers as well as the set-up, significance, or validity of the research projects and findings. This means that most studies are either very broad, touching a variety of backgrounds and aspects, or extremely narrow, limited to very specific projects or unique combinations of settings. Furthermore, the various interests and backgrounds also shape the methods and theoretical frameworks that are used in research, creating a big diversity in relation to these aspects as well. As a result, existing studies can rarely be compared to each other and identifying relevant research and appropriate methods and frameworks is a challenging task. The literature read for this thesis provided many more leads and hints that might have been interesting to follow up. This includes for example literature from crowdsourcing projects (like Wikipedia or open source software development) that might provide further insights on moderation in virtual collaborative projects. However, covering all the possible related aspects would have been an impossible task, since both citizen science and online forum moderation can be connected to such a large number of concerns and backgrounds.

Another challenge was the decision how to categorize the moderator posts. While good examples of pedagogical, social, and technical support were found in the datasets, involved participation and interaction seemed to occur less often and moving forward and managerial support were difficult to identify. Letting categories emerge from the data rather than trying to fit them into an existing categorization of a different background (educational setting) might have been a more appropriate approach. The 'roles' identified in an interview study (Frith, 2014) might be even more difficult to apply for categorization, since they were not meant as a framework for categorization and present the view of moderators rather than users or observers. However, combined together the two sets of distinguishing features did offer some insights about the datasets. For example, instead of the quite limited notions of interaction support and managerial introduced by Asterhan (2011), Frith offers the idea of moderators as 'information architects' that shape the overall setup of the forums rather than

individual interactions and as ‘tone-setters’, which might include a combination of interaction support, moving forward, and social support. Maybe a distinct way of categorizing moderator actions specifically in citizen science projects that includes aspects from education and other settings of moderation might be necessary instead of trying to fit the data into existing categorizations. Nevertheless, the existing distinctions might offer valuable points of departure, much like the theoretical framework of a ‘community of purpose’ (Wiggins & Crowston, 2010) can be contextualized as lying between a learners’ community and a community of interests.

Finally, the selection of the two citizen science projects and the specific data analyzed present limitations for this study and lead to opportunities for further research. The selection of posts by the time of a specific moderators’ activity in Foldit and by selecting a subforum in Galaxy Zoo was a very simple way of narrowing down the large amounts of data from the discussion forums. This means that only a small amount of the data collected was actually used for analysis. Other subforums in Galaxy Zoo or other moderators’ posts in Foldit might yield additional or differing results, especially in the distribution of the categories identified. Furthermore, the data was only categorized by one person; comparing categorizations by several raters could have helped to eliminate possible errors or ambiguities. Additionally, the big differences of the two citizen science projects and the use of different data selection criteria for practical reasons make it very difficult to compare the results from the two discussion forums. Choosing data by the same criteria or data from more similar projects (e.g. two different Zooniverse projects) might make the findings from the two forums more comparable to each other. It should also be noted that the datasets included only what is visible to the public in one particular channel of communication for each project. Galaxy Zoo and Foldit both include other channels of communication, some of which are publicly accessible (e.g. wiki, blog) and some of which are visible only to specific users or groups of users (e.g. personal messages, chat). The collected data from the two forums include only a small part of each projects’ overall community. Participants might use other communication channels, only read but not contribute to discussions, or even choose not to interact with other participants at all, even if they are active contributors to the project tasks. Furthermore, the two projects studied, Galaxy Zoo and Foldit, both are virtual citizen science projects using only online environments, whereas many other citizen science projects include physical interactions with the environment (Wiggins & Crowston, 2011). Therefore, other projects will probably show very different communities and patterns of interaction. Online discussion forums might not play an important role, or not be implemented at all. This means that the findings from Galaxy Zoo and Foldit may not be generalizable to other citizen science projects, especially not to smaller or more localized projects where participants might interact in face-to-face discussions rather than online. Nevertheless, even when online discussion moderators might not play less important roles, the acts of mediating between stakeholders, recognizing users’ contributions, encouraging learning, and curating knowledge from the project participants might be important in other citizen science projects as well. The considerations about the generalizability of the findings to the entire project or to other projects lead to the question, if and how the moderator roles identified here might be taken over by other actors in different communication channels or project setups, which might be an interesting topic for further research.

## Summary and Conclusion

The aim of this thesis was to explore how moderators act in the discussion forums of two citizen science projects, Foldit and Galaxy Zoo. The article presented in the main part of this thesis explains how data from both forums was selected and analyzed using a categorization from work on moderation in an educational synchronous discussion environment (Asterhan, 2011). Possible moderator roles resulting from interviews with technical help forum moderators (Frith, 2014) were also considered; the observations on these roles were summarized in the article and a detailed explanation of examples was given in the first part of the thesis. Resulting from these categorizations, a discussion about the roles of moderators observed in the dataset was presented. Although there are differences between the two forums and even between different moderators, four major moderator roles in citizen science projects emerged from the data: moderators can help to connect different stakeholders, collect and organize information from different sources, encourage and motivate participants by acknowledging their contributions, and offer users support in learning. Additionally, it was discussed how discussion forums in general, and moderator activities in particular, can support and encourage learning, the latter being frequently stated as a goal or possible benefit for participants in citizen science projects.

In trying to connect citizen science and forum moderation, this study draws from a range of different backgrounds. That made identifying and choosing appropriate literature, methods and theoretical frameworks very challenging. However, the finished research can contribute to those various backgrounds, offering new observations on citizen science as well as forum moderation. This study considers forum moderation in discussion forums that lie in between the scope of existing literature, i.e. educational and public discussion forums. Considering the different theoretical frameworks that might be used to conceptualize the citizen science projects as a community revealed that the applied framework by Wiggins and Crowston (2010) also seems to fall in this gap between the formal educational 'learners' communities' and the much looser connected 'communities of interest'.

The particular focus of activities and roles varies between different moderators. The observations made in this study might be characteristic only for virtual citizen science projects or even unique to Galaxy Zoo or Foldit. However, the roles identified seem to concern important aspects of citizen science projects. Analyzing which actors can or do fulfill these functions might give valuable insights even for other citizen science projects.

## References

- Alonso, A. D., O', M., & Shea, N. A. (2012). Moderating virtual sport consumer forums: exploring the role of the volunteer moderator. *International Journal of Networking and Virtual Organisations*, 11(2), 173. <http://doi.org/10.1504/IJNVO.2012.048332>
- Arnt, A., & Zilberstein, S. (2003). Learning to perform moderation in online forums. In *Proceedings IEEE/WIC International Conference on Web Intelligence (WI 2003)* (pp. 637–641). IEEE Comput. Soc. <http://doi.org/10.1109/WI.2003.1241285>
- Asterhan, C. S. C. (2011). Assessing e-moderation behavior from synchronous discussion protocols with a multi-dimensional methodology. *Computers in Human Behavior*, 27(1), 449–458. <http://doi.org/10.1016/j.chb.2010.09.008>
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, 59(11), 977–984. <http://doi.org/10.1525/bio.2009.59.11.9>
- Brace-Govan, J. (2003). A method to track discussion forum activity: The Moderators' Assessment Matrix. *The Internet and Higher Education*, 6(4), 303–325. <http://doi.org/10.1016/j.iheduc.2003.08.003>
- Brossard, D., Lewenstein, B., & Bonney, R. (2005). Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education*, 27(9), 1099–1121. <http://doi.org/10.1080/09500690500069483>
- Cerulo, L., & Distanto, D. (2013). Topic-driven semi-automatic reorganization of online discussion forums: A case study in an e-learning context. In *2013 IEEE Global Engineering Education Conference (EDUCON)* (pp. 303–310). IEEE. <http://doi.org/10.1109/EduCon.2013.6530121>
- Chou, C. C. (2002). A comparative content analysis of student interaction in synchronous and asynchronous learning networks. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (Vol. 2002-Janua, pp. 1795–1803). IEEE Comput. Soc. <http://doi.org/10.1109/HICSS.2002.994093>
- Convery, I., & Cox, D. (2012). A review of research ethics in internet-based research. *Practitioner Research in Higher Education*, 6(1), 50–57. Retrieved from <http://194.81.189.19/ojs/index.php/prhe/article/view/100>
- Crall, a. W., Jordan, R., Holfelder, K., Newman, G. J., Graham, J., & Waller, D. M. (2012). The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. *Public Understanding of Science*, 22(6), 745–64. <http://doi.org/10.1177/0963662511434894>
- Danchak, M. M., & Kenyon, K. (2002). Threaded discussion as a tool in the asynchronous technology classroom. *Proceedings - Frontiers in Education Conference*, 1, T1E/14–T1E/18. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-0036957183&partnerID=tZOtx3y1>
- De Wever, B., Keer, H. Van, Schellens, T., & Valcke, M. (2010). Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior*, 26(4), 516–523. <http://doi.org/10.1016/j.chb.2009.08.008>

- De Wever, B., Van Winckel, M., & Valcke, M. (2008). Discussing Patient Management Online: The Impact of Roles on Knowledge Construction for Students Interning at the Paediatric Ward. *Advances in Health Sciences Education*, 13(1), 25–42. <http://doi.org/10.1007/s10459-006-9022-6>
- Delort, J.-Y., Arunasalam, B., & Paris, C. (2011). Automatic Moderation of Online Discussion Sites. *International Journal of Electronic Commerce*, 15(3), 9–30. <http://doi.org/10.2753/JEC1086-4415150302>
- Dickinson, J. L., Shirk, J., Bonter, D., Bonney, R., Crain, R. L., Martin, J., ... Purcell, K. (2012). The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment*, 10(6), 291–297. <http://doi.org/10.1890/110236>
- Dong, F. (2012). Controlling the internet in China: The real story. *Convergence: The International Journal of Research into New Media Technologies*, 18(4), 403–425. <http://doi.org/10.1177/1354856512439500>
- Edwards, R. (2014). Citizen science and lifelong learning, 46(2), 132–145.
- Engeström, Y. (2007). From Communities of Practice To Mycorrhizae. *Communities of Practice: Critical Perspectives.*, 1–20. <http://doi.org/10.4324/9780203101339>
- Franzoni, C., & Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. *Research Policy*, 43(1), 1–20. <http://doi.org/10.1016/j.respol.2013.07.005>
- Frith, J. (2014). Forum moderation as technical communication: The social web and employment opportunities for technical communicators. *Technical Communication*, 61(3), 173–184. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84907032223&partnerID=tZOtx3y1>
- Galaxy Zoo. (n.d.). The Story So Far. Retrieved from <http://www.galaxyzoo.org/#/story>
- Gee, J. P., & Hayes, E. (2012). Nurturing affinity spaces and game-based learning. In Constance Steinkuehler, Kurt Squire, & Sasha Barab, Eds., *Games, Learning, and Society: Learning and Meaning in the Digital Age*. (pp. 129–155).
- Ghosh, A., Kale, S., & McAfee, P. (2011). Who moderates the moderators? In *Proceedings of the 12th ACM conference on Electronic commerce - EC '11* (p. 167). New York, New York, USA: ACM Press. <http://doi.org/10.1145/1993574.1993599>
- Ghosh, A., Kale, S., & McAfee, P. (2011). Who Moderates the Moderators? Crowdsourcing Abuse Detection in User-Generated Content Categories. *Proceedings of the 12th ACM Conference on Electronic Commerce*, 167–176. <http://doi.org/10.1145/1993574.1993599>
- Guldberg, K., & Pilkington, R. (2007). Tutor roles in facilitating reflection on practice through online discussion. *Educational Technology and Society*, 10(1), 61–72. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-33947668793&partnerID=tZOtx3y1>
- Henri, F., & Pudelko, B. (2003). Understanding and analysing activity and learning in virtual communities. *Journal of Computer Assisted Learning*, 19(4), 474–487. <http://doi.org/10.1046/j.0266-4909.2003.00051.x>
- Herman, S. (2010). Career HOPES: An Internet-delivered career development intervention. *Computers in Human Behavior*, 26(3), 339–344. <http://doi.org/10.1016/j.chb.2009.11.003>

- Hsieh, Y.-H., & Tsai, C.-C. (2012). The effect of moderator's facilitative strategies on online synchronous discussions. *Computers in Human Behavior*, 28(5), 1708–1716. <http://doi.org/10.1016/j.chb.2012.04.010>
- Jackson, C., Østerlund, C., Crowston, K., Mugar, G., & Hassman, K. D. (2015). Motivations for Sustained Participation in Citizen Science: Case Studies on the Role of Talk. *48th Hawaii International Conference on System Sciences*, 1624–1634. <http://doi.org/10.1109/HICSS.2015.196>
- Jenkins, E. W. (1999). School science, citizenship and the public understanding of science. *International Journal of Science Education*, 21(7), 703–710. <http://doi.org/10.1080/095006999290363>
- Jones, K. M. L., Stephens, M., Branch-Mueller, J., & de Groot, J. (2016). Community of practice or affinity space: A case study of a professional development MOOC. *Education for Information*, 32(1), 101–119. <http://doi.org/10.3233/EFI-150965>
- Jordan, R. C., Ballard, H. L., & Phillips, T. B. (2012). Key issues and new approaches for evaluating citizen-science learning outcomes. *Frontiers in Ecology and the Environment*, 10(6), 307–309. <http://doi.org/10.1890/110280>
- Jordan, R. C., Gray, S. A., Howe, D. V., Brooks, W. R., & Ehrenfeld, J. G. (2011). Knowledge Gain and Behavioral Change in Citizen-Science Programs. *Conservation Biology*, 25(6), 1148–1154. <http://doi.org/10.1111/j.1523-1739.2011.01745.x>
- Jyothi, S., McAvinia, C., & Keating, J. (2012). A visualisation tool to aid exploration of students' interactions in asynchronous online communication. *Computers & Education*, 58(1), 30–42. <http://doi.org/10.1016/j.compedu.2011.08.026>
- Kienle, A., & Ritterskamp, C. (2007). Facilitating asynchronous discussions in learning communities: the impact of moderation strategies. *Behaviour & Information Technology*, 26(1), 73–80. <http://doi.org/10.1080/01449290600811594>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. *Learning in doing* (Vol. 95). Retrieved from <http://books.google.com/books?id=CAVIOrW3vYAC&pgis=1>
- Luczak-Rösch, M., Tinati, R., Simperl, E., Kleek, M. Van, Shadbolt, N., Simpson, R., ... Simperl, E. (2014). Why Won't Aliens Talk to Us? Content and Community Dynamics in Online Citizen Science. *Proceedings of the Eighth International Conference on Weblogs and Social Media, {ICWSM} 2014, Ann Arbor, Michigan, USA, June 1-4, 2014*. Retrieved from <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/download/8092/8136> <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8092>
- Masters, K., Oh, E. Y., Cox, J., Simmons, B., Lintott, C., Graham, G., ... Holmes, K. (2016). Science learning via participation in online citizen science, 15(03).
- Raddick, M., Bracey, G., & Gay, P. (2010). Galaxy zoo: Exploring the motivations of citizen science volunteers. *Astronomy Education Review*, 9(1). Retrieved from <http://portico.org/stable?au=pgg3ztfdp8z>
- Roth, W. M., & Lee, S. (2004). Science Education as/for Participation in the Community. *Science Education*, 88(2), 263–291. <http://doi.org/10.1002/sce.10113>

- Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., ... Jacobs, D. (2012). Dynamic Changes in Motivation in Collaborative Citizen-Science Projects. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12* (pp. 217–226). <http://doi.org/10.1145/2145204.2145238>
- Savio, L. Del, Prainsack, B., & Buyx, A. (2016). Crowdsourcing the Human Gut . Is crowdsourcing also „ citizen science ”?, *15*(03), 1–16.
- Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution*, *24*(9), 467–71. <http://doi.org/10.1016/j.tree.2009.03.017>
- Thomas, J. (2013). Exploring the use of asynchronous online discussion in health care education: A literature review. *Computers & Education*, *69*, 199–215. <http://doi.org/10.1016/j.compedu.2013.07.005>
- Trumbull, D. J., Bonney, R., Bascom, D., & Cabral, A. (2000). Thinking Scientifically during Participation in a Citizen-Science Project. *Inc. Sci Ed*, *84*, 265–275. [http://doi.org/10.1002/\(sici\)1098-237x\(200003\)84:2<265::aid-sce7>3.0.co;2-5](http://doi.org/10.1002/(sici)1098-237x(200003)84:2<265::aid-sce7>3.0.co;2-5)
- Vlachopoulos, P., & Cowan, J. (2010a). Choices of approaches in e-moderation: Conclusions from a grounded theory study. *Active Learning in Higher Education*, *11*(3), 213–224. <http://doi.org/10.1177/1469787410379684>
- Vlachopoulos, P., & Cowan, J. (2010b). Reconceptualising moderation in asynchronous online discussions using grounded theory. *Distance Education*, *31*(1), 23–36. <http://doi.org/10.1080/01587911003724611>
- Wals, A. E. J., Brody, M., Dillon, J., & Stevenson, R. B. (2014). Convergence between science and environmental education. *Science*, *344*(May), 583–584.
- Wang, J., Li, Q., & Chen, Y. P. (2010). User comments for news recommendation in social media. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval - SIGIR '10* (p. 881). New York, New York, USA: ACM Press. <http://doi.org/10.1145/1835449.1835663>
- Wiggins, A., & Crowston, K. (2010). Developing a conceptual model of virtual organisations for citizen science. *International Journal of Organisational Design and Engineering*, *1*(1), 148–162. <http://doi.org/10.1504/IJODE.2010.035191>
- Wiggins, A., & Crowston, K. (2011). From conservation to crowdsourcing: A typology of citizen science. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 1–10. <http://doi.org/10.1109/HICSS.2011.207>
- Wise, A. F., Saghafian, M., & Padmanabhan, P. (2012). Towards more precise design guidance: Specifying and testing the functions of assigned student roles in online discussions. *Educational Technology Research and Development*, *60*(1), 55–82. <http://doi.org/10.1007/s11423-011-9212-7>
- Xie, K., Yu, C., & Bradshaw, A. C. (2014). Impacts of role assignment and participation in asynchronous discussions in college-level online classes. *Internet and Higher Education*, *20*, 10–19. <http://doi.org/10.1016/j.iheduc.2013.09.003>
- Zhang, W., Cao, X., & Tran, M. N. (2013). The structural features and the deliberative quality of online discussions. *Telematics and Informatics*, *30*(2), 74–86. <http://doi.org/10.1016/j.tele.2012.06.001>



## Appendix 1