Which Way Did They Go? Newcomer Movement through the Zooniverse

Corey Brian Jackson, Carsten Østerlund, Veronica Maidel*, Kevin Crowston, Gabriel Mugar
School of Information Studies, Syracuse University
*Zooniverse, Adler Planetarium
{cjacks04,costerlu,crowston,gmugar}@syr.edu and *veronica@zooniverse.org

ABSTRACT
Research on newcomer roles in peer production sites (e.g., Wikipedia) is characterized by a broad and relatively well-articulated set of functionally and culturally recognizable roles. But not all communities come with well-defined roles that newcomers can aspire to occupy. The present study explores activity clusters newcomers create when faced with few recognizable roles to fill and limited access to other participants' work that serves as an exemplar. Drawing on a mixed method research design, we present findings from an analysis of 1,687 newcomers' sessions in an online citizen science project. Our analysis revealed three major findings: (1) newcomers' activities exists across six session types; (2) newcomers toggle between light work sessions and more involved types of production or community engagement; (3) high-level contributors contribute large volumes of work but comment very little and another group contributes large volumes of comments, but works very little. The former group draws heavily on posts contributed by the latter group. Identifying shifts and regularities in contribution facilitate improved mechanisms for engaging participants and for the design of online citizen science communities.

Author Keywords
Online Communities; Crowdsourcing; Citizen Science

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Peer production and crowdsourcing communities (online production communities collectively) rely on participants to contribute work and take on responsibilities to manage and sustain the community. To be successful, communities need to maintain a critical mass of active participants [4, 22]. For most communities, maintaining a critical mass requires a flow of active and engaged newcomers into the community to replace those who drop out. Reflecting the importance of the topic, a significant body of research has emerged on newcomers in peer-production communities. Wikipedia studies hold a prominent position [2, 3, 12, 22, 30] together with FLOSS investigations [12, 32]. These studies largely agree on the diversity of participation patterns [3, 22]. Many newcomers show up briefly and contribute little before they exit. But a few newcomers almost as quickly become core contributors to the community.

A number of studies have examined how newcomers learn to be productive in communities. One form of guidance comes from roles, i.e., organizational formations exhibiting structural and functional commonalities in behavior patterns that are culturally recognizable by organizational members [3]. A number of online communities exhibit well-defined roles that define how members spend time contributing. For example, Arazy et al. [3] described twelve Wikipedia roles (e.g., registered user, technical administrator). They organized these roles into a power hierarchy with unregistered users at the bottom and, at the top, the Benevolent Dictator, founder Jimmy Wales. Likewise, many FLOSS teams have structurally and culturally defined roles. Most projects have an onion-like structure with a few core participants at the center that have the rights to commit code [33, 34] and a large periphery of occasional contributors. Over time, peripheral members may become core members by making valued contributions.

The content of roles may be explicitly defined in a few cases, but more often, knowledge about roles is passed on through experience or contact with core users. For example, Dejean and Jullien [11] reported that sites like Wikipedia attract participants with some prior knowledge about the forms of engagement required. They found that many newcomers have been mentored by existing Wikipedians in their social networks before they joined the community. In FLOSS projects, core members provide examples of expected contribution in their own work. Other studies emphasized participants' access to practice [1, 6], feedback from experienced members [15, 18, 24], and relationship-building activities [12].

However, not all online production communities come with clear roles that newcomers can aspire to occupy. Zooniverse projects, for instance, like many other citizen science projects, explicitly distinguish only the roles of participant,

moderator, science team member, and platform developer. Of these four roles, newcomers can only realistically aspire to become participants or, less commonly, moderators. The other roles are restricted to those with a prior relationship to the project. A newcomer to the Galaxy Zoo project, for example, is unlikely to join the science team alongside its professional astronomers.

Furthermore, not all communities provide access to members' experiences from which roles can be learned vicariously. The lack of visibility may be inadvertent or deliberate. For instance, for methodological reasons, Zooniverse projects do not allow participants to observe other participants' data coding practices so that each participant's coding of scientific data is done without influence from others. Even when observation is possible, online communities often do not have enough journeymen members to adequately provide feedback and support the relationship building that is heralded as important in peer-production communities.

While communities that lack defined roles or opportunities to learn via observation may seem odd or dysfunctional, this situation describes most newly-formed communities in which the organizational structure (roles included) is still emerging, and exemplary work is scarce. Studying such communities is thus revelatory for understanding these kinds of organizations and provides an opportunity to examine how the self-organizing actions of newcomers structure the community.

A further difficulty in understanding newcomer behavior is finding an appropriate unit of analysis for comparison. Comparing entire contribution histories of community participants is complicated by the great diversity of these histories. As noted above, many newcomers only "test the waters" before dropping out, and so have very short histories, while a few go on to be active contributors with long and diverse histories. It is difficult to make meaningful comparisons of such diverse histories, beyond noting the difference in their duration. To address this issue, in this paper we compare contributions grouped into sessions, where a session is defined as a set of contributions made close together in time (i.e., within a few minutes) and separated from other contributions by a larger gap (e.g., hours). The research questions guiding this study are:

1. Given the lack of formal roles to guide activity, what clusters of activity do we find among newcomers?
2. What behaviors patterns emerge across newcomer sessions at the individual level?
3. Why do newcomers contribute with these behavioral patterns?

RELATED WORK

The literature on newcomers in online production communities has tended to focus on the individual participant as the unit of analysis. The research findings can be summarized along three dimensions. First, researchers often divide the community into active, less active, and passive participants. Second, some studies have found little change in newcomers' behaviors after their initial engagement with a project, while others found significant change in the type of roles participants eventually engage in. Third, among the studies finding changes in participation, some described these movements as rather sequential while others discovered non–sequential progressions. We address each of these dimensions in turn.

Active, less active, and passive participants: A number of prior studies divided participants into active members who contribute the majority of work, less active participants, and passive members who take advantage of the benefits offered without contributing themselves to community activities [22, 27, 29]. Crowston and Fagnot [10] divided participants into non–contributors, initial contributors, continuing contributors, and meta–contributors. In Wikipedia, Priedhorsky et al. [31] found that the top 0.1% of all editors, by number of edits, performed 44% of the work. In online social networks (e.g., Facebook and LinkedIn) Benevenuto et al. [5] concluded that passive participants (i.e., lurkers) constituted 92% of all users. For example, Nonnecke and Preece [25] examined health–support and software–support discussion lists and found that lurkers accounted for 46% and 82% of the participants, respectively. Muller et al. [24] showed that 72.2% of users are lurkers in an enterprise file–sharing service. Citizen science is no different; Eveleigh et al. [13] reported that a small core (6%) contributes the vast majority of the work (85%). Some scholars have characterized the less active citizen scientists as educational lurkers [1, 25, 29]. Mugar et al. [23] observed that newcomers often start learning about a site by navigating the site, reading "FAQ" or "About" pages, and doing some light annotation work.

Core–periphery structure captures these different levels of engagement, where a small number of participants carry a large burden of the work and responsibility at the core, and the majority of contributors engage in few tasks at the community's periphery [3,6,7,10]. The distinction between core and periphery tends to rely on quantitative measures such as the number of contributions, views of content, visits, time spent on site, and interactions with other members [3, 22]. Qualitative studies commonly solicit participants' experiences of member satisfaction, community inclusion, and the quality of interaction.

Change or no change: While the literature concurs on big–small and core–periphery distinctions, one finds no consensus on participants' movement between these different levels of engagement and what such a journey involves. One can broadly divide the literature into studies that find little change in the types of behavior and level of engagement exhibited by a newcomer after their initial participation and studies depicting newcomers as moving through a number of stages leading towards deeper engagement.

First, a number of studies suggested that newcomers from their very first contribution exhibit lasting behavioral traits, such that we can know from the beginning whether a participant's future engagement will be big or small [11]. In these studies, little significant change occurs in activity levels over time, and active participants behave differently from others. Among Wikipedians and Cyclopath contributors, Panciera et al. [27, 28] found that power editors are "born not made".

On their first day of editing, the top 2.5% of editors contributed more edits per editor than the rest of the registered editors [27, 28]. These high performers maintained a constant level of participation for the majority of their tenure with Wikipedia. Fang and Neufeld [14] found similar patterns in open source software development.

In contrast, other research has found that newcomers do change behavior over time. Typically, this movement takes place along the periphery–core axis where newcomers start out at the periphery and gradually move towards the center, involving a great diversity of activities and responsibility. Many of the stage models build on the notion of legitimate peripheral participation (LPP) [21, 35], suggesting that newcomers start out as legitimate but peripheral participants who learn through observing and doing simple tasks. Gradually, they engage in more complex activities and gain access to core community practices. Comparably, Preece and Shneiderman [30] synthesized prior work in this area into a successive ladder of activities and roles, termed the "reader to leader" framework. Newcomers start out as consumers of a site, e.g., reading content on Wikipedia, and gradually work their way up the ladder as they become more active. A fraction of those readers start contributing by editing or engaging in other core activities. A portion of those contributors begins to work with others to develop and maintain resources and become collaborators. At the final stage, leaders emerge among the collaborators as they begin to work on governance and policy development. Likewise, Crowston and Fagnot [10] described a movement from initial participation to sustained contribution and eventually meta–contribution. The latter group may no longer contribute much to a project's core task[s], but dedicate most of their time to project administration and maintenance, an equally necessary task for the project to be successful.

Sequential or non–sequential: Stage models tend to suggest a sequential move from the periphery towards the core. Newcomers move in an ordered progression from readers, for example, to contributors, collaborators, and then leaders. However, a couple of studies seemed to suggest that not all participants follow a sequential path. Gray [17] argued that some participants toggle between the center and periphery and that sequential and non–sequential progressions both constitute legitimate learning patterns in online communities. Likewise, in a study of leadership roles in Wikipedia, Arazy et al. [3] found no unidirectional movements between center and core. Rather, Wikipedians tended to take a non–sequential path, where some newcomers move straight from entry level activities to the community's core, while others transition from the core back to the community's periphery. And sometimes leaders simply retire from key positions.

Returning to the broader question about what activity clusters newcomers create when faced with few recognizable roles to fill and limited access to other participants' work, the literature leaves us with a number of specific questions. First, what range of activity levels do we find among newcomers? Second, do we find change in newcomers' participation, as they become members of an online community? Third, if change takes place, does it involve a sequential move from the periphery towards the core or not? The existing literature offer few hints as to what patterns of activity are likely to emerge at the session level and whether these amass into specific behavioral patterns at the level of the individual participant in situations largely absent of formal roles. There are many theories concerning motivation to contribute to projects. However, motivation is usually considered for participation in general: few theories distinguish motivations for different kinds of activities available on a site.

METHODOLOGY

To answer our research questions, we analyzed data from the Planet Hunters citizen science project. Our study uses a mixed methods sequential exploratory design (SED) in which data from different sources are collected and analyzed separately, but embedded during synthesis [9]. Our SED relied on both quantitative and qualitative data from the same project, though not from the same respondents. The qualitative study (Study 1) was conducted to understand how participants make sense of the project, their role, and which resources they use to support their participation. Nine months after Study 1 concluded, changes to the site allowed us to track pages participants viewed as they navigated with site. Page view data was collected for two months. We then combined page view data with server logs which contained the total contribution history of participants. We analyzed the data to identify trends in page views and contribution to the project (Study 2). Finally, we interpreted the data through triangulation of both data sources with participant names anonymized. The intent of combining these two sources of data is to suggest what beliefs and attitudes might be behind the observed behavior.

Research Setting

Planet Hunters is an online astronomy citizen science project in which astronomers seek the help of volunteers to filter data collected from the Kepler space telescope. The task requires volunteers make annotations on light curve images (periodograms) to record the apparent presence of transits (i.e., dips in the light curve which suggesting the presence of a planet). As of May 2015, approximately 300,000 volunteers have annotated more than 20 million periodograms.

One notable feature of this setting is that the system only shows the periodogram and its metadata during annotation. The markings of other participants are not visible so as to reduce annotation bias. However, once an annotation is submitted, the system asks participants if they would like to discuss their work with other participants. During this time, participants can read the comments of other members or post a comment.

Planet Hunters Activities

Beyond annotating periodograms, participants can perform other activities that support the project. For example, participants can post comments on project namespaces (e.g., discussion boards, Talk pages), provide feedback to peers, start discussion topics explaining the results of their independent

research, or build collections of transits. Each activity supplements project resources (including scientists), in that additional training materials are made known, valuable feedback provided and independent analysis discussed.

While not all participants post comments, many read content which is equally valuable in supporting the project and individual participant goals. Since the annotation task can be ambiguous and feedback isn't always provided, participants (newcomers in particular) can read discussions of more experienced participants and come to understand why and how annotations are made. It follows that, as participants read and learn more, the quality of their annotation improves. As participants navigate the site, they might find additional materials which help support their participation or become involved in activities that provide additional motivation for participation.

/object: After submitting an annotation, participants can read and post comments to conversations about the image they annotated. Conversations about work practices and annotation consensus can be observed. Participants can leave hashtags (#) linking conversations with similar topics.

/discussion: Linked in the navigation bar, pages contain general astronomy topics (e.g., glitches, astronomy topics, links to astronomy research). Participants can start discussions and contribute to existing conversations. Some participants post findings based on independent analysis to "publish" work and solicit input from other members.

/group: Participants can append images on one page which supports independent analysis.

/collection: Participants append images they find useful or interesting; pages are public and frequently built around potential transits.

/other: Homepages and user profile pages.

Study 1: Qualitative Study

Data collection began in July 2012 and lasted for 1 year and 6 months. The data collected consist of 15 semi–structured interviews, 2 focus groups, and 5 diary entries. Participants were recruited through email, solicitations on Planet Hunters, and fliers posted in across our campus. The goal of these interviews was to understand learning and motivation in citizen science projects. Interviews with project participants focused on how they learn to annotated and how they interacted with the social spaces of the project. Interviews with project administrators focused on the organizational aspects of the project, such as availability of learning resources, project governance, and interactions between the scientists and participants. Interviews were recorded, transcribed, and interviewer notes were composed. Two focus group sessions helped us gain insight into resources used to support early participant learning. We asked focus group participants to create an account, complete the tutorial, and begin working in the project. After each participant worked for 45 minutes, all participants came together and discussed difficulties they had with annotating and how they addressed those challenges. Lastly, three doctoral students engaged in participant observation, a form of virtual ethnography [19], which emphasizes

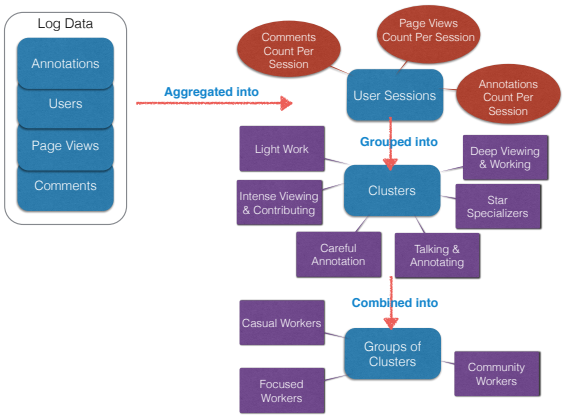

Figure 1. Trace data processing steps.

participation in the online environment by the researcher. As participant observers, we created accounts, completed tutorials, consulted project resources, annotated images, and interacted in social spaces. In analyzing data, practice–oriented theories [8, 26] framed our understanding of how newcomers engage in PH, and guided our analysis of transcribed interviews, field notes, diary entries, and contributions to social spaces.

Study 2: Quantitative Study

Data collection began in September 2014 and concluded October 2014. The purpose of this study was to understand when participants used project resources (e.g., FAQ Pages, tutorials, comments left by other participants) and if the frequency of use changed overtime. We analyzed the data using trace ethnography [16], which highlights the history of participant activity as it appears in server logs. Figure 1 shows how the data collected in Study 2 was processed. First, we aggregated page view data and server logs containing annotations and comments of each participant. Second, we aggregated data by activity at the session level to achieve counts of activity per session. A session is defined as a submission of an annotation where no more than 30 minutes exists between the current and next annotation. We then computed additional data points about each session which included number of page views prior to starting a session, (2) number of annotations submitted during a session, and (3) average time spent in a session.

Third, sessions were grouped based on counts of dimensions (e.g., number of contributions to /object, /discussion, annotations) using a k-means clustering algorithm. Because variance is high, data was standardized prior to clustering. The within-groups sum of squares plot (i.e.,"elbow" plot) revealed seven was the optimal number of clusters. K-means was run with the above-mentioned features 500 times with 500 different starting points, and the best fit was chosen based on the minimum within-cluster sum of squares. To confirm that the k-means clustering was identified distinctive clusters, a MANOVA analysis with session dimensions as dependent variables and the identified clusters as the independent variable was performed. Fourth, sessions were combined for each

participant to reveal a history of participation. Lastly, each history of participation was grouped and the groups named.

RESULTS

Findings from Study 1

From the qualitative data, we identified three themes related to newcomer behavior. First, participants arrive at the project with varying backgrounds in relation to astronomy. One interviewee, Robert, a project moderator, is employed as a writer for a tech magazine and spends his spare time doing astrophotography. Robert spoke of his participation as a newcomer and his contribution of more than 3,000 annotations in the first month of his membership in Planet Hunters (PH). Jill, a newcomer, who has a background in physics, chemistry, and biosciences but no formal background in astronomy, contributed more than 1,000 annotations. Other interviewees spoke of their educational and hobbyist backgrounds in astronomy which motivated them to explore the project. Many reported owning telescopes, subscribing to astronomy publications, and having a lifelong interest in astronomy.

Second, interviewees spoke of different modes of participation. Some interviewees described their participation as contributing to the science goals of the project by filtering data for professionals. Other interviewees spoke of a desire to learn everything they could and frequently consulted project resources like the tutorial and /discussion pages in order to make accurate annotations. Justin, another newcomer, reflected on his experiences downloading metadata about a star, an activity typically performed by experienced members, in order to perform independent analysis. Unfortunately, the data was beyond his grasp, and he concluded he would need more advanced training for the advanced analysis.

Third, participants spoke of different levels of engagement with social spaces. Some interviewees expressed hesitation to post comments. April, for example, a newcomer with only two hours of project tenure, stated, "*I don't personally appreciate it.*" and when asked about feedback, responded, "*Not really, I'm only trying to help.*" Other newcomer interviewees were unaware of the social spaces or, because of barriers, failed to participate. Michael, a newcomer, spoke of a "*highly technical vocabulary*" which limited his participation. Another newcomer, Lisa, stated she perceived talk (/object page) as a way to communicate unusual or novel features in the periodogram and provide feedback to the project. Lisa feels she is not experienced enough to discover anything unusual or novel and therefore doesn't share in /object pages. While posting in social spaces seems unachievable for many newcomers, experienced members and moderators note the significance of viewing social spaces. Roger, for example, often referred to the tutorial, other people's comments on /object pages, and blog posts that described different characteristics in periodograms relevant to project goals.

The result of our qualitative findings demonstrates participants take on different roles and patterns of behavior. For example, given varying backgrounds, it seems that newcomers engage in different activities with varying regularities. Also, since experienced participants and moderators describe pages like /object as beneficial for participant learning, we were curious if many newcomers engaged in these spaces early on to support their learning.

	Views	Contributions (No. Participants)
/collection	329	16 (14)
/discussion	184	35 (21)
/group	477	4 (4)
/object	7,773	967 (277)
/other	882	0
Annotations		63,252

Table 1. Description of newcomers activity (views and posts created). The number in parentheses in the Contributions column represents the number of users contributing to the activity.

Findings from Study 2

We now turn to the log data collected in Study 2. A total of 1,684 newcomers created accounts during our data collection period. Newcomers contributed to the project by posting and viewing pages. Table 1 shows the aggregated activities for this group of newcomers. The distribution of activity counts resemble that for other online communities, in which many participants contribute irregularly or in small amounts, and a handful contribute the majority of the content. Newcomers contributed across 2,978 sessions, individually contributing across an average of 1.7 (sd. = 2.7) sessions. Figure 2 shows the distribution of sessions and reveals 71% of newcomers (N = 1,203) dropped out of the project after one session and only 2% (N = 34) contributed across seven or more sessions. The newcomers contributed a total of 63,252 annotations (mean = 37.54, sd. = 169.62), visited 9,652 pages (mean = 5.7, sd. = 35.62), and posted 1,023 comments (mean = .61, sd. = 4.7).

Observing Science and Community

Newcomers viewed /object pages frequently, navigating to these pages on 7,773 occasions. Visits to /object pages are closely associated with annotation in that comments are tied to specific periodograms revealing the context of what the annotators observed. Examples of comments are: "*Transits Grey Rock looking 1.0005-14/ 1.0002-11 Grey Transit 1.0007-11. One Dark And Round Planet Looking at area 1.0002-08*" and "*#Transit seen day 27*". Comments on /object pages display different levels of work practice and specificity around annotations. In the former example, cues point to where the annotator observed a transit and presents additional characteristics like color, while the latter points to the presence of transit and its location but without additional description. Still, both may be valuable for newcomers as they provide indications about how to do the annotation task.

In addition to substantive comments, newcomers can also observe the comments of other newcomers (some frequently identify themselves as novices), like one participant who stated, "*I am new to this. What is the explanation for this pattern? The granular detail doesn't appear to be as deep for this as the others were.*" This newcomer questions identifying patterns in the periodogram and received a response: "[username] *-possible transits would indeed be below normal light level. 1'st step is to look for dots that show periodicity*"

[1]. It is likely that many newcomers have similar questions about identifying transits, and comments similar to these offer an example of work practices that serve as educational opportunities for newcomers as they learn to identify transits in periodogram images.

Viewing pages supports the findings in [23] that suggest that early participation is spent visiting /object (or Talk pages) in an attempt to observe work practices of experienced citizen scientists. The /collection, /discussion, and /group pages attracted less attention than did the /object pages. In these project pages, we find more scientific analysis, through which more experienced members are coordinating work. For example, in the /discussion space a post between experienced members titled, "*How to fold a light curve [periodograms] using Excel*" [2] provides, over a series of posts, a tutorial on conducting analyses using Microsoft Excel. A /collection page titled "*Known Eclipsing Binary*" [3] provides a collection of periodograms that newcomers can observe to learn the visual characteristics of eclipsing binaries on a periodogram.

Newcomer Contributions

Newcomers submitted 63,252 annotations to help scientists identify transits. Contribution volume varied among newcomers as is evident by the large variance (mean = 37.54, sd. = 169.62). Contribution to the other pages was infrequent among newcomers (mean = .61, sd. = 4.7) and as with viewing behaviors, data was skewed towards contributions to /object pages where 277 (16%) newcomers posted comments. Posting on /object pages is where newcomers seek feedback from more experienced participants. One participant stated, "*Hi everyone, I am a new guy in the planthunter. I would like to know your opinion about the transit of the object APHF10180918 shown in Fig.*" or "*Awesome find. I just started can you look at my findings?*" to receive feedback. Interestingly, viewing comments before posting allows newcomers to observe the proper form (i.e., specifying the the day, location) for conveying the presence of transits. In general, posts to /object pages allow newcomers to discuss their work with other participants, receive feedback, and engage in other conversations with community members. We find other newcomers seeking similar feedback in /object pages, and we suspect this is the primary motivation for contribution to /object pages. We did not observe much activity beyond /object pages. Only a small numbers of newcomers contributed to /discussion (N = 21), /collection (N = 14), and /group (N = 4). These namespaces represent areas where advanced scientific discussion or independent science work by other citizen scientist occur.

RQ1: Cluster Signatures

The k-means cluster analysis yielded seven clusters of activity patterns from the 2,978 user sessions. We removed one cluster from the further discussion because it included only a single session that was characterized by extreme values in most dimensions. Conversely, one cluster in our analysis contained 90% (N = 2,690) of the sessions. Table 2, ordered by median session number in the cluster, shows the mean of activities performed for each cluster. The quantitative signatures of each cluster allow us to describe dominant activity patterns. The results of the MANOVA analysis confirm that there are statistically significant differences between the clusters on the sessions dimensions. $F(60, 13864.302) = 5328555227, p<.0005$; Wilks' $\lambda = .000$; partial $\eta^2 = 1$.

Light Work: The majority of sessions were in this category. Session activities are dominated by annotation work (mean = 17.23) and rarely include viewing or posting to other namespaces. The median session number for a session in this cluster is 1, meaning that the first session of many newcomers falls in this category. The small number of activities suggests dabbling behavior in annotation to test out the project before making long-term commitment.

Intense Viewing/Contributing: The median session number for this type of session is again 1, meaning that a few newcomers began their careers in this cluster. This cluster is characterized by high volume of annotations (mean = 210.5) and views in the /object (mean = 55.6) and other (mean = 21) namespaces. Surprisingly, sessions characteristics show early viewing activity in the /discussion namespace (mean = 6), suggesting some newcomers immerse themselves (through observation) in advanced scientific discussion. The differences between this cluster and the Light Work cluster are in viewing behavior and volume of contribution to annotation. These differences are interesting in that they are indicative of interest in the project beyond dabbling.

Careful Annotation: This cluster comprises 109 sessions where newcomers were engaged in careful annotation work. The cluster is distinguished by the average time spent in the session per annotation (mean = 7.5 minutes). An interesting characteristic of these sessions is the frequency of viewing and contributing to the /object namespace (mean = 3.87 and 1.3 respectively), while submitting only 5.67 annotations. It seems that participants are paying close attention to annotation work, and this potentially reveals a practice of combing through comments on /object namespace to learn from other members and asking questions about their work.

Talking & Annotating: Sessions in this cluster exhibit the full gambit of work in project annotation, contribution, and viewing in namespaces. This cluster shares similarities to the previous two clusters, but differs in the volume of comments contributed to the /object namespace (mean = 21.3), more than eight times as many as in the other clusters. These values suggests sessions where participants annotate images and then post a comment (asking questions or describing work). These sessions also showed a noticeably larger number of namespace views prior to starting session work (mean = 7.17), suggesting participants arrived at the site to engage in reading /other pages, check their user profiles, or view /object or /discussion pages before starting annotation work.

Deep Viewing & Working: The second largest session cluster reveals intense focus on contributing annotations and viewing namespaces. These session activities are characterized by participation in the full range of viewing and contributing be-

[1] http://oldtalk.planethunters.org/objects/APH73009062
[2] http://oldtalk.planethunters.org/discussions/DPH100suj9
[3] http://oldtalk.planethunters.org/collections/CPHS0002gh

		Light Work	Intense Viewing & Contributing	Careful Annotating	Talking & Annotating	Deep Viewing & Working	Star Specializers
Views (μ)	/collection	0.03	3.2	0.18	0.16	1.06	0.5
	/discussion	0	6.27	0.1	0.11	0.35	0.75
	/group	0.06	2.73	0.17	0.42	1.68	3.5
	/object	1.21	55.6	3.87	44.79	16.69	5.25
	/other	0.08	21.2	0.35	0.32	2.22	1.25
Contributions (μ)	/collection	0	0.07	0	0	0.08	0
	/discussion	0	0.33	0.01	0	0.06	0
	/group	0	0	0	0	0	1
	/object	0.09	2.53	1.3	21.32	0.82	0.75
	Annotations	17.23	210.53	5.67	33.84	87.98	10.75
Descriptive	Views Before (μ)	0.19	0.27	0.83	2.47	7.17	1
	Time Spent (μ)	0.84	0.93	7.5	3.49	0.92	1.61
	Median Session	1	1	2	2	3	4
	Cluster Size	2690	15	109	19	140	4

Table 2. Mean cluster values ordered by median session number for cluster observations.

haviors (except on /group pages). Sessions in this cluster have the second largest number of annotations (mean = 87.98) and viewed pages spanning the project. These sessions also revealed interesting numbers of namespace views (mean = 7.2) prior to starting annotation work.

Star Specializers: This session cluster, though comprised of just a few sessions, shows increased participation in one of the project's scientific spaces (i.e., /group). In these sessions, little annotation work was submitted, but participants instead moved their participation towards a space where participant–driven annotation, querying, and analysis takes place (e.g., group and discussion pages). While this cluster includes only four sessions, contribution to /group pages, where multiple annotated periodograms are gathered, represents contribution to advanced science work by synthesizing how a transit spans multiple observations.

Our clustering reveals combinations of activity. Sessions types exhibit varying levels of engagement with project activities where some newcomers are focused on annotation work (i.e., Light Work), others intensely focus on viewing /object spaces and making annotations (i.e., Intense Viewing & Contributing), and another group engages the community by posting and responding to comments (i.e., Talking & Annotating).

RQ2: Behavioral Patterns

To derive participation histories from the the k-means clustering, we combined and ordered user sessions in a bar chart (Figure 2). Figure 2 shows session timelines for newcomers. The x–axis represents the session number with unique participants on the y–axis. Each bar represents a session and is colored coded based on the session cluster type identified by the k–means clustering and named by the researchers. For example, the first session of the last user in Figure 2 was Careful Annotation and the second session was Talking & Annotating. For all users in Figure 2, we identified similarities in contribution timelines to help distinguish different participation types. After visually inspecting user session histories, we observed a few session clusters. First, many newcomers remained in a single session type (433 newcomers with more than one session had all the same type of session). Second, we observed contribution patterns where newcomers oscillate between three clusters types: (1) Deep Viewing & Working and Light Work, (2) Intense Viewing & Contributing and Light Work, and (3) Talking & Annotating and Light Work.

RQ2 & RQ3: Inferring Newcomer Movement

At the individual level, we identified four patterns of newcomers based on their work. First, as is typical of peer production communities, many individuals dropped out after only one session. In our dataset, 1203 (71%) participants dropped out. However, the work that dropouts contribute is not insignificant. Dropouts contributed 17,615 annotations (mean = 14.5) in their single session. We suspect participants who dropped out were testing the project to determine if they would continue membership. Of the individuals who dropped out, 1152 (95%) had a Light Work session.

The remainder of our analysis is centered on session activities of 481 newcomers who continued in the project. We identified three patterns of contribution based on groupings of participation histories: (1) Casual Workers, (2) Community Workers, and (3) Focused Workers.

Casual Workers

The work necessary to support the science goals of the project is submitting annotations. We labeled as Casual Workers those participants who contributed only in sessions described as Light Work (this pattern is not shown in Figure 2). The total number of Casual Workers was 378. Casual Workers contributed across many sessions (mean = 3), submitted 22,484 annotations, and posted 93 (mean = .08) comments, but only had Light Work sessions.

User 4 (Figure 3), for example, contributed to over 25 sessions, but remained in Light Work throughout his membership. User 4 began viewing many namespaces early in his membership though pageviews seemed to decline in volume after the tenth session.

Interviews provided insight into this behavior. For instance, newcomer Lie Yao reported having no prior experience in astronomy and being unaware of the social spaces (i.e., /collections, /objects, or /discussion). The social side of the project was not important to Yao, who stated, "*That is not what I do...other people are talking about what they do on Facebook. I just do the marking*". Surprisingly, Yao's view on feedback and social interaction contradicts literature about newcomer socialization, which suggests feedback is important to sustain

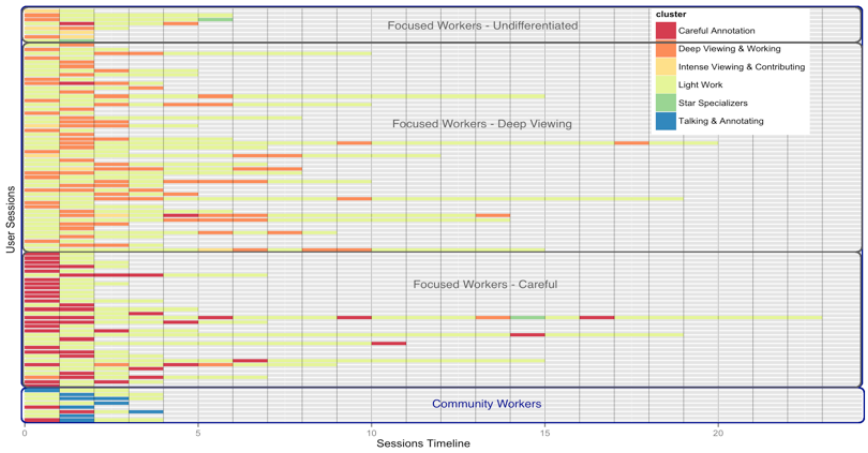

Figure 2. Illustration of the shifts in session cluster membership. Participants with fewer than two sessions and participants not shifting sessions during their participation (N= 1584) are not shown. The session history of two newcomers is not represented in the figure; these participants had more than 40 sessions making it difficult to visualize the remaining (N= 97) participants.

Figure 3. Illustration of the shifts in session clusters for a sample of participants. The first chart shows annotation history, the second shows page views, and the last shows contributions. Each string of points represent a single participant. The size of the point corresponds to the normalized volume of activity and the color reveals which cluster for that session.

newcomer participation. Yao states, "*Feedback is not that important. I'm always trying my hardest*".

Even when newcomers were aware of the social spaces, participants described apprehension to contribute. Ariel, a newcomer with two months' tenure at the time of our interview, revealed that contributing to /object and /discussion pages would be challenging, stating:

"*I don't know [if] I've done [anything] significant to warrant discussion. They ask do you want to discuss this star and I always press no because hell I don't know what to discuss about it...it seems like to me we've got a bunch of astrophysicists sitting around discussing these things and that sure isn't my expertise.*"

User 4's log data along with Yao's and Ariel's interview revealed that some newcomers are quite content performing routine annotation work isolated from the social setting, though we also find potential barriers to further contribution.

Community Workers

We describe as Community Workers the newcomers who post comments early in their participation. Though only eight newcomers were identified as Community Workers, they contributed 618 annotations (mean = 23) and posted 216 comments (mean = 8). The activity by this small group of newcomers manifests itself in posting comments for the purpose of asking questions, justifying their about work practices, providing feedback to other individuals, and organizing work by making use of hashtags (e.g., # transit) to retrieve periodograms in the future. Newcomers who engage in these activities provide important functional roles that are not apparent prior to joining the project. Their activities help further curate data, organize project resources, and foster the social connections amongst other participants.

User 23 (Figure 3), for example, contributed across 40 sessions, and his third session activities were categorized as Talking & Annotating. User 23's early sessions are dominated viewing and making comments to PH's social spaces. Analyzing traces of User 23s annotations, views, and contributions (i.e., comments) reveals he spent more time investigating the project by viewing pages than making annotations. In sessions 3 thru 6, User 23 visited more pages than he submitted annotations. During these sessions User 23 posted between 14 and 30, abnormal behavior for the majority of new-

comers. In later sessions, User 23's behaviors reveal what appears to be a relationship between annotations and contribution, that is, every annotation leads to comment.

In our qualitative interviews, we heard accounts of newcomers participating in the social spaces of PH. Contribution comes in the form of asking questions about their work, seeking feedback. Newcomers post comments like, "*This is my first Transit at 4.5? Hope I am on the right track*" or from another newcomer, " *I've done some searching for an answer but have found only generic guides to variability. My question is this: I see that the y-scales are dependent on the light measurements.*" The former shows a newcomer seeking feedback about whether they found an annotation in the graph located at day 4.5 on the periodogram, while the latter seeks clarification on whether they have correctly identified the purpose of the scaling on the image. While drawn from different datasets, we suspect User 23 and the newcomers mentioned above take on similar roles providing useful content for the community and newcomers in particular to learn form. These comments provide valuable insight into the work practices (e.g., identifying transits or providing justification for work) and typical issues (e.g., glitches) that might arise.

Focused Workers

The final pattern of behavior we identified was newcomers who view namespaces and contribute project work but still pay little attention to the social side of the project. We identified 89 newcomers who we labeled as Focused Workers. These newcomers contributed 16,562 annotations (mean = 34) but posted comments infrequently (N = 191, mean = 0.4), suggesting a focus on contributing to the science goals of the project, that is, annotation. These workers are distinguished from Casual Workers by their engagement in sessions other than Light Work. When we visually inspected Figure 2, we noticed two patterns of movement between clusters; first, oscillation between Deep Viewing & Working and Light Work, and second, oscillation between Careful Annotation and Light Work.

Deep Viewing & Working and Light Work: These pairings represent the behavior of toggling between viewing and making annotations. User 31 (Figure 3), for example, contributed across 15 sessions. Her session history reveals many annotations and views, but very few comments. User 30 contributed frequently (68 sessions) and maintained a history of oscillating between Light Work and Deep Viewing. Inspecting trace data for User 30 revealed a pattern of viewing pages prior to beginning a session. We suspect, User 30 would review interesting conversations in /discussion and on /object pages to absorb knowledge about astronomy or the project. Interestingly, User 30 viewed 1,681 pages, but posted only 11 comments throughout their history.

Qualitative interview data again shed light on the behavior patters of users 30 and 31. Similar to User 31, Annie, a newcomer interviewee, revealed she has submitted fewer than 50 annotations and has never posted a comment in the project. In her work on annotating images, Annie describes that she moved back and forth between the annotation interface and the help pages of the project whenever she was unsure about her work, stating:

"*Well when you start looking over the images you can always have a click back on the help button, so you can have a few images where you know what you're doing and then you'll have one that will bring up something different and then you can always go back and really go through some of the quick tutorials then you can understand what you're looking at.*"

In this example, we find a newcomer who is a low volume classifier, taking her time to check in with the help features of the site to figure out how to do the work.

Careful Annotation and Light Work: Some participants regularly classified, but averaged more time working through annotations. User 23's (Figure 3) recent session history is identified as Careful Annotation. Although beginning with Community Work, User 23's later history reveals a focus on individual work rather than contribution to social spaces. User 23 frequently spent many minutes annotating periodograms. For example, in the 20 sessions identified as Careful Classifying, User 23 spent a minute analyzing the periodogram.

In interviews, we find Emily, a newcomer who was motivated by a lifelong love for astronomy. Emily described how she was in "learning mode" as a newcomer, going to the talk page of every periodogram she classified in the hope of learning from comments that other users posted.

"*I was looking at the comments like after every [annotation] that you do, you can click afterwards discuss and see what people say about this particular target. So I was just in learning mode, you know just thinking okay, I want to do this but I want to do this right.*"

In addition to searching for learning opportunities, Emily took annotating very seriously, describing how the task was "grueling" and that she would take an hour to classify some stars, stating:

"*I'm not the kind of person who just does something and it doesn't matter. There's a seriousness to what you're doing, and I wanted to be serious about it and if I was going to do that and spend so much time doing it, I wanted to do it right.*"

Similarly, Henry, an experienced participant, describes how, after discovering the zoom tool in the annotation interface, he changed his process of annotation from a quick scan to a more thorough analysis, spending upwards of 20 minutes per periodogram in hopes of finding transiting planets. The story of participants like User 23, Emily, and Henry reveal some participants are dedicated to accurately marking the periodograms. Henry's discovery of the zoom feature for example and could provide possible explanation about why some participants take more time to work.

Undifferentiated Workers: Finally, a few newcomer session histories did not exhibit any recognizable patterns and thus did not fall into the two prior categories. These participants either had no dominant activity (i.e., a tie in cluster counts that comprised their history) or had a session timeline with an equal number of session cluster types.

DISCUSSION

There are a significant number of recent studies of newcomers in peer production sites such as Wikipedia and FLOSS projects. These communities tend to offer well-defined functionally and culturally recognizable roles that newcomers can aspire to occupy. However, the literature offers few clues regarding how newcomers fare as they engage with crowdsourcing communities that have few roles that are either formally defined or modeled by more senior participants. Our study of newcomers' activity clusters in the first few months of their engagement with Zooniverse offers new insights to the debate by combining analysis at the session and individual participant level.

Production

The literature divides participants into active, less active, and passive members. In contrast, we find ta diversity of participation when we consider activity at the session level as compared to the individual participant level. Careful Annotation sessions contribute, on average, 5.67 annotations compared to 210.5 for Intense Viewing & Contribution sessions (Table 2). When making comments to /objects pages, Light Work sessions are less frequent contributors (mean= 0.09) compared to Talking & Annotating sessions (mean = 21.3).

At the individual level, 1,203 of the newcomers dropped out after only one session in which they did (on average) a few annotations and almost no Object contributions. Another group of newcomers (totaling 379) became what we labeled Casual Workers, who contributed an average of 20 annotations and 0.08 Objects over the course of 2.5 months. In the same period, eight Community Workers made 24 annotation and 9 Object contributions on average. The Focused Workers, numbering 89, contributed with an average of 34 annotation and 11.4 Object discussions. Yet, it should be noted that we do find a range among individuals within each of these types of workers.

Intriguingly, newcomers engaging in more than one type of session cluster toggled between Light Work and other activity clusters (i.e., Focused Workers and Community Workers). In fact, Light Work appears to serve as a baseline session type. After one or more intense sessions newcomers always return to a bit of Light Work, breezing through a few annotations with little involvement with other system features and activities. However, newcomers engaged in Focused Work do not jump around among session types. With few exceptions, this group tended to toggle between two clusters of activities, one being Light Work. For instance, we see people move between Deep Viewing Working and Light Work or Careful Annotation and Light Work. The same nonlinear pattern does not seem to fit newcomers engaging in Community Work. This very small group switches between three or more session types including either Talking & Annotating or Star Specializers and some combination of Light Work, Careful Annotation, and Deep Viewing & Working.

Learning

While Zooniverse offers newcomers few structurally and culturally recognizable roles to choose among, our findings suggest that steady activity patterns do emerge. However, our analysis cannot confirm any clear change in participation over the course of the study. For instance, we do not see a clear trajectory in newcomers' participation as suggested by LPP where newcomers start out engaged in Deep Viewing & Working as they learn from more experienced participants and later contribute to the community through Talking & Annotating. Instead, newcomers seem to stick to one type of session, as in the case of Casual Workers, or two in the case of Focused Workers, and more when it comes to Community Workers. This does not mean that we do not see indications of learning among some participants. In particular, Focused Workers tend to spend a lot of time on each annotation, repeatedly checking what other people are talking about or discussing the object they have annotated. Others browse the project pages before they start annotating and regularly check /object, /collections, and /group pages while working. These activity patterns may represent learning behaviors where newcomers set out to learn as much as they can about the project and the images they annotate. Often concerned about getting it right, this small population of newcomers dedicate themselves to careful work.

Talk

Our findings showed that a small number of newcomers become Community Workers contributing to Talk (i.e., /object) or delving into user-generated research taking place around the Group (i.e., /group) space and Discussion (i.e., /discussion) pages. While only a small number of participants contribute to Talk, most newcomers view discussions of periodograms through the Talk feature and Focused Workers spend a lot of time studying other people's discussion of periodograms in Talk. Even Casual Workers view periodograms in the Talk pages. In a narrow sense this finding may not be surprising given results from peer production research indicating that just a few active participants contribute most content [11]. Past research also suggest that active contributors often exhibit these characteristics within their first few sessions [11, 27, 28]. These findings also seem to corroborate research findings that top performers in peer production communities are born, not made [27]. This being said, our findings highlight not only the importance of Talk and Discussion features for online production communities as spaces for learning, but also the significance of the relationship between Focused Workers and Community Workers. Community Workers provide learning material (i.e., useful comments) and Focused Workers consume materials which help determine whether their work is accurate. One might expect that a project would benefit if more newcomers learned to contribute to Talk and other communal features. That might offer a richer ecosystem, benefiting both current and future newcomers.

Implications for Online Communities

From quantitative work and supplemental qualitative interviews with experienced participants and community managers, it is apparent that newcomers possess varying levels of commitment to the project, prior exposure to astronomy, and inclination for self-directed learning. If in fact the members of the crowd are born unequal, there are consequences

for system feature design and community management. Absent formal roles in many citizen science projects, roles beyond scientist, system developer, and moderator, making the community work visible to newcomers is important. In many online communities there are activities beyond the stated goal of the project that are essential to sustaining the community but are not formally recognized (e.g., in the case of Planet Hunters contributing to discussion as well as to annotation of periodogram images). In another study of Planet Hunters, Jackson et al. [20] noted citizen scientists are motivated by providing additional benefits to the community. Engaged participants often seek out their own niche. One participant, Patrick, specialized in finding glitches that others might have mistaken as transiting planets. This type of community work is important in order to address possible confusion of citizen scientists and to prevent wasting valuable time on unimportant discussion. While "glitch identifier" isn't a role, and probably never will become formally recognized in Planet Hunters, it gives purpose to participants like Patrick and offers an important if unrecognized service to the community.

LIMITATIONS AND FUTURE WORK

The study has a number of limitations. First, we gathered data for only a two and one half month period for Study 2 (Quantitative Study), ending when the Planet Hunters project website was upgraded to a new version of the software. While the study focuses on newcomers, future research would benefit from a significantly longer time horizon, which would allow us to track how specific activity clusters change over time.

Second, while the populations from Study 1 and Study 2 didn't overlap, the interviews helped us develop a qualitative sense of newcomer's perception of Planet Hunters, their use of specific project features, and learning over time. However, the timing of these interviews did not allow us to validate the session clusters emerging from the quantitative analysis. In future research we plan to select interview subjects based on their activity clusters identified through quantitative analysis.

Nevertheless, combining analysis at the session and individual participant level appears promising. All of our qualitative work suggests that more people experience the system session by session and often share their experiences about the process they use to do work. Further analysis at the session level might offer interesting new insights that could help develop a better sense of when people are likely to stop working, for example. As with physical exercise, one could hypothesize that people start a session with some lighter work before they dive into the heavy lifting, followed by some less strenuous activities before they stop.

Finally, the study research would benefit from comparisons across multiple citizen science projects. For example, the OldWeather project has a small core of participants and a small population of newcomers, and their participation might look a bit different than the population we focused on. We are currently tracking newcomers across other Zooniverse projects and hope to compare the clusters emerging from those contexts with the emerging results from Planet Hunters.

The category of Focused Workers also calls for additional analysis. We notice that this class of newcomer spans a number of session types, from Intense Viewing & Contribution to Careful Annotation. Further analysis may suggest helpful distinctions in this group.

ACKNOWLEDGEMENTS

We thank the millions of Zooniverse participants and the Zooniverse team for access to data. This material is based on work supported by the National Science Foundation under Grant No. IIS 12-11071.

REFERENCES

1. Judd Antin and Coye Cheshire. 2010. Readers are not free-riders: reading as a form of participation on wikipedia. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. ACM, 127–130.
2. Judd Antin, Coye Cheshire, and Oded Nov. 2012. Technology-mediated contributions: editing behaviors among new wikipedians. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 373–382.
3. Ofer Arazy, Felipe Ortega, Oded Nov, Lisa Yeo, and Adam Balila. 2015. Functional Roles and Career Paths in Wikipedia. In *the 18th ACM Conference*. ACM Press, New York, New York, USA, 1092–1105.
4. A Bechmann and S Lomborg. 2013. Mapping actor roles in social media: Different perspectives on value creation in theories of user participation. *New Media & Society* (2013).
5. Fabrício Benevenuto, Tiago Rodrigues, Meeyoung Cha, and Virgílio Almeida. 2009. Characterizing user behavior in online social networks. In *IMC '09: Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*. ACM Request Permissions, New York, New York, USA, 49–62.
6. Susan L Bryant, Andrea Forte, and Amy S Bruckman. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. ACM, 1–10.
7. Brian Butler, Elisabeth Joyce, and Jacqueline Pike. 2008. Don't look now, but we've created a bureaucracy. In *Proceeding of the twenty-sixth annual CHI conference*. ACM Press, New York, New York, USA, 1101–10.
8. Karin Knorr Cetina, Theodore R Schatzki, and Eike von Savigny. 2005. *The Practice Turn in Contemporary Theory*. Routledge.
9. John Creswell. 2013. Research design: Qualitative, quantitative, and mixed methods approaches.
10. Kevin Crowston and Isabelle Fagnot. 2008. The motivational arc of massive virtual collaboration. In *Proceedings of the IFIP WG 9.5 Working Conference on Virtuality and Society: Massive Virtual Communities*.

11. Sylvain Dejean and Nicolas Jullien. 2014. Big From the Beginning. Assessing Online Contributors' Behavior by Their First Contribution. *Assessing Online Contributors' Behavior by Their First Contribution.(February 7, 2014)* (2014).
12. Nicolas Ducheneaut. 2005. Socialization in an Open Source Software Community: A Socio-Technical Analysis. *Computer Supported Cooperative Work (CSCW)* 14, 4 (July 2005), 323–368.
13. Alexandra Eveleigh, Charlene Jennett, Ann Blandford, Philip Brohan, and Anna L Cox. 2014. Designing for dabblers and deterring drop-outs in citizen science. In *the 32nd annual ACM conference*. ACM Press, New York, New York, USA, 2985–2994.
14. Yulin Fang and Derrick Neufeld. 2009. Understanding Sustained Participation in Open Source Software Projects. *Journal of Management Information Systems* 25, 4 (April 2009), 9–50.
15. R Stuart Geiger, Aaron Halfaker, Maryana Pinchuk, and Steven Walling. 2012. Defense Mechanism or Socialization Tactic? Improving. In *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*.
16. R Stuart Geiger and David Ribes. 2011. Trace ethnography: Following coordination through documentary practices. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on*. IEEE, 1–10.
17. Elizabeth Louise Gray. 2004. Informal Learning in an Online Community of Practice. *Journal of Distance Education* 19, 1 (2004), 20–35.
18. Aaron Halfaker, Aniket Kittur, and John Riedl. 2011. Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work. In *Proceedings of the 7th international symposium on wikis and open collaboration*. ACM, 163–172.
19. Christine Hine. 2000. *Virtual Ethnography*. Sage Publications, Thousand Oaks.
20. Corey Brian Jackson, Carsten Østerlund, Gabriel Mugar, Katie DeVries Hassman, and Kevin Crowston. 2015. Motivations for Sustained Participation in Crowdsourcing: Case Studies of Citizen Science on the Role of Talk. In *HICSS '15: Proceedings of the 2015 48th Hawaii International Conference on System Sciences*. IEEE Computer Society, 1624–1634.
21. Jean Lave and Etienne Wenger. 1991. *Situated Learning. Legitimate Peripheral Participation*. Cambridge University Press., NY.
22. Sanna Malinen. 2015. Understanding user participation in online communities: A systematic literature review of empirical studies. *Computers in Human Behavior* 46, C (May 2015), 228–238.
23. Gabriel Mugar, Carsten Østerlund, Katie DeVries Hassman, Kevin Crowston, and Corey Brian Jackson. 2014. Planet hunters and seafloor explorers: legitimate peripheral participation through practice proxies in online citizen science. In *the 17th ACM conference*. ACM Press, New York, New York, USA, 109–119.
24. M J Muller, D R Milien, and J Feinberg. 2009. Information curators in an enterprise file-sharing service. *ECSCW 2009* (2009).
25. Blair Nonnecke and Jennifer Preece. 2000. Lurker demographics: counting the silent. In *CHI '00: Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM Request Permissions, New York, New York, USA, 73–80.
26. C Østerlund and P Carlile. 2005. Relations in practice: Sorting through practice theories on knowledge sharing in complex organizations. *The Information Society* (2005).
27. Katherine Panciera, Aaron Halfaker, and Loren G Terveen. 2009. Wikipedians are born, not made. In *Proceedinfs of the ACM 2009 international conference*. ACM Press, New York, New York, USA, 51.
28. Katherine Panciera, Reid Priedhorsky, Thomas Erickson, and Loren G Terveen. 2010. Lurking? cyclopaths?: a quantitative lifecycle analysis of user behavior in a geowiki. In *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Request Permissions, New York, New York, USA, 1917–1926.
29. Jennifer Preece, Blair Nonnecke, and Dorine Andrews. 2004. The top five reasons for lurking: improving community experiences for everyone. *Computers in Human Behavior* 20, 2 (March 2004), 201–223.
30. Jennifer Preece and Ben Shneiderman. 2009. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction* 1, 1 (March 2009), 13–32.
31. Reid Priedhorsky, Jilin Chen, Shyong Tony K Lam, Katherine Panciera, Loren G Terveen, and John Riedl. 2007. Creating, destroying, and restoring value in wikipedia. In *GROUP '07: Proceedings of the 2007 international ACM conference on Supporting group work*. ACM Request Permissions, New York, New York, USA, 259–268.
32. Yixin Qiu, Katherine J Stewart, and Kathryn M Bartol. 2010. Joining and Socialization in Open Source Women's Groups: an Exploratory Study of KDE-Women. . In *Open Source Software New Horizons*, Pär J gerfalk, Cornelia Boldyreff, Jesús M González-Barahona, Gregory R Madey, and John Noll (Eds.). Springer, 438.
33. Eric S Raymond. 2007. The cathedral and the bazaar. *First Monday* 3, 2 (Feb. 2007).
34. Peter Wayner. 2000. *Free for All*. Harperbusiness.
35. Etienne Wenger. 2010. Communities of Practice and Social Learning Systems. *Organization* 7, 225 (2010).