Bring on Board New Enthusiasts!
A Case Study of Impact of Wikipedia Art + Feminism
Edit-a-thon Events on Newcomers

Rosta Farzan, Saiph Savage, and Claudia Flores Saviaga

[1] University of Pittsburgh
[2] West Virginia University
[3] Universidad Nacional Autonoma de Mexico (UNAM)
rfarzan@pitt.edu
saiph.savage@mail.wvu.edu
saviaga@gmail.com

Abstract. Success of online production communities such as Wikipedia highly relies on a continuous stream of newcomers to replace the inevitable high turnover and to bring on board new sources of ideas and labor. However, these communities have been struggling with attracting newcomers, especially from a diverse population of users. In this work, we conducted a case study on how organizing an offline co-located event over a short period of time contributes to involving newcomers in the online community. We present results of our multiple-source quantitative analysis of Wikipedia Art+Feminism edit-a-thon as a case of such events. The results of our analysis shows that such offline events are successful in attracting a large number of newcomers; however, retention of the newcomers stays as a challenge.

1 Introduction

Online production communities such as Wikipedia have been enjoying omnipresent success stories; however, the success stories are accompanied by significant challenges. An important challenge identified within a range of online production communities is ensuring a stream of newcomers to replace an inevitable high turnover they face and to attract sources of new ideas and new labor [18]. This problem is even more intensified when such communities try to recruit and retain newcomers from a more diverse population of users. For example, Wikipedia not only has faced a plateaued growth of new editors over the recent years [28], it has been particularly struggling with the challenge of attracting female editors [17].

In response to the challenge of attracting new members and developing commitment, a number of online production communities such as open source software communities and Wikipedia have tried to organize offline co-located gatherings to foster recruitment and integration of newcomers. The importance and occurrence of offline interactions in conjunction with online communications have been acknowledged by a number of studies [2], [29], [15], [21],[12]. A few studies have attempted to assess and quantify the impact of offline gatherings on online participation. While there is strong

evidence in support of supplementing online interactions with offline gatherings, there is also evidence that highlights the challenges arising as a result of offline connections. In some cases, offline connections can lead to weakening of the online interactions as a result of creating stronger clicks among those who can meet offline, or by shifting the interactions offline thereby reducing online interactions [25]. It has been argued that offline gatherings can promote stronger bonds that lead to stronger bonding social capital, but this is accomplished at the expense of decreasing weak ties and bridging social capital [27]. More recently, multiple case studies of open source software hackathons presents results on how the structure of such events influences the outcomes; especially in terms of advancing the production goals of the community as well as building social ties [31], [30].

In this work, we are specifically interested to investigate how offline co-located gatherings affect participation of newcomers in online production communities. Hackathons-like events often have been viewed as onboarding programs. It has been argued that the intense training and social bonding opportunities provided in such offline events can particularly benefit newcomers and socialize them to the community by teaching them the performing and social rules [30]. However, no prior work has particularly investigated the impact of these offline gatherings on newcomers' socialization into the online community and the mechanism with which they can influence newcomers. As more resources and times are dedicated to these collective efforts, it becomes critical to understand the impact such events have on online production communities.

2   Socialization of newcomers in online communities

Prior research has investigated socialization of newcomers in a variety of online communities, including open-source software (e.g. [9]), social media (e.g. [20]), and peer-production (e.g. [10]). They have particularly compared and contrasted socialization approaches in traditional organizations with strategies being employed online. While the results are not conclusive, often they have documented that many online communities lack specific strategies for socializing newcomers into the community [6]. An important factor identified by a number of studies as positively influencing newcomers' commitment to the online community, is interaction between newcomers and existing members [1, 20, 4]. Newcomers who receive feedback and communication from the existing members of the community, even if the feedback is criticism, are more likely to continue participating in the community.

Several investigations have focused on creating computational systems and methods to increase volunteer workforces [3], especially for political causes [24]. Other systems bootstrap off social media to access large pools of people to facilitate the recruitment process. Brady et al. [3] showed that it was feasible to recruit volunteers from people's Facebook friends to help the blind. Savage et al. [24] showed the potential of using online bots to recruit people from Twitter to do micro-volunteering for a cause. Nevertheless, these platforms are at present incomplete and have mostly focused on recruitment of newcomers and not longer term commitment of newcomers [26]. While these approach can bring an influx of volunteers to a collaborative effort, they rarely

maintain the volunteers long-term [13]. This is especially because the recruitment is often not followed by any longer-term engagement mechanism.

Other approaches have focused precisely on creating work flows that encourage long-term engagement of volunteers. Such platforms have sometimes sandboxes where newly recruited volunteers can have personalized and detailed feedback on their work from experts [22]. The sandboxes let newly engaged volunteers to become integrated into the cause under a friendly welcoming environment. This can help in their retention. However, the effectiveness of such approaches on socialization and retention of newcomers has not been researched. Other approaches have engaged new crowds of volunteers with simple lightweight feedback processes [7]. These approaches showcase how new volunteers can be retained through lightweight guided contribution. However, there seems to be a lack of connection between specially approaches In this work we take a look at newcomer socialization processes that are offline and take consequently more time from both longer term community members and the newcomers. We analyze and contrast such processes with these other methods to engage newcomers.

3 Wikipedia Art+Feminism Edit-A-Thon

Wikipedia has often been named as one of the most successful examples of online production communities and product of collective wisdom. Despite its enormous success, it has also been facing a great deal of challenges over time. In particular, as highlighted by researchers at Wikimedia foundation, attracting newcomers is one of the key challenges faced by the Wikipedia community [11]. While some argue that committed members of the community exhibit different behavior and signs of commitment from early on [23], it has been shown that active strategies employed by the community and by the newcomers, such as friendly interactions with experienced members [22], active socialization approaches within Wikipedia [6], constructive feedback and avoiding undermining of their goodwill efforts [14, 32] can increase the likelihood of future commitment.

The challenge of attracting newcomers becomes even more demanding when trying to target a more diverse population and those who have been underrepresented in the existing community. At the same time, it has been shown that diversity can play an important role in success of production communities such as Wikipedia [5]. Since 2011, several studies have highlighted a phenomenon of gender imbalance in Wikipedia that indicates only around 15% of Wikipedia contributors identify as female and a very small percentages of Wikipedia contributions are made by female Wikipedians [8, 19]. It has further been documented that this gender inequality has resulted in quantitative and qualitative inequality in representation of topics more attractive to female readers as well as inequality in representation of biographies of notable women[19]. This inequality happens despite the fact that women are generally more likely to participate in volunteer and community based activities and they are more likely to participate in social sites such as Facebook [19].

In response to this challenge, since 2014, a group of Wikipedia and feminism enthusiasts have been organizing Wikipedia Art+Feminism edit-a-thon events [4]. Edit-a-thon events are collocated all-day events bringing together novices and experienced

---

[4] `https://en.wikipedia.org/wiki/Wikipedia:Meetup/ArtAndFeminism`

Wikipedia editors. The goal of these events is to increase the coverage of female representation in Wikipedia and to encourage female editorship. The events are advertised on the Web and through various social media platforms [5] [6]. It is particularly highlighted in the advertisement of the event that no prior editing experience is required and as one of the first activity of the day, a tutorial on editing Wikipedia is presented to the participants. By 2015, 75 Art+Feminism edit-a-thon events had been organized which attracted 1,500 participants and resulted in creating or improving 900 Wikipedia articles. In the current work, we focus on investigating the impact of the most recent set of edit-a-thon events, organized in 2016, on attraction and retention of newcomers into Wikipedia. We hope that our research can provide insights to organizers of such events to better understand how these event contribute to their goals. We also hope that by studying these events we can better understand the process of integrating more minorities into the production process.

4 Research questions

We argue that an event such as edit-a-thon can influence newcomers' socialization process through two different mechanisms: (1) the focused gathering of an edit-a-thon event can provide newcomers with intense training opportunities to learn how to get work done in Wikipedia that can lead into more effective and consequently higher level of participation; (2) the collocated gatherings can build connections between newcomers and experienced Wikipedians helping them to build strong identification with Wikipedia that can lead to higher level of commitment and participation. Therefore, we have formulated the following research questions in better understanding of the impact of edit-a-thons on newcomers:

RQ1: how does attending an edit-a-thon event influence subsequent participation of newcomers on Wikipedia?

RQ2: does attending an edit-a-thon lead to bond and connections among participants?

RQ3: how do production and social interaction factors influence the retention of edit-a-thon newcomers in editing Wikipedia articles?

5 Research methods and data collection

To address our research questions, we conducted quantitative analysis on archival log data available on Wikipedia and Twitter. Using crawling approaches, Wikipedia API[7], and Twitter API[8], we collected data related to 59 edit-a-thon events happening in the US in the period of Jan 2016 until March 2016.

To study our first research question, we collected information on newcomers' logged behavior on Wikipedia during and after the edit-a-thon events to assess at what level

---

[5] `http://art.plusfeminism.org/`

[6] `https://www.facebook.com/events/876331705807795/`

[7] `https://www.mediawiki.org/wiki/API:Main_page`

[8] `https://dev.twitter.com/rest/public`

they participated on the event day and subsequently after the event. To better understand their level of commitment, we attempted to contrast subsequent Wikipedia editing participation of newcomers who attended an edit-a-thon event with comparable newcomers who joined Wikipedia independent of edit-a-thon events. We identified specific editing tasks that newcomers performed on the day of the event, including creating user pages, editing in Sandbox pages, and editing article pages. Based on our experience with Wikipedia and Wikipedians community, we classify each of these editing activities as representing different familiarity and identification with Wikipedia. We collected information on newcomers' activities with respect to each of these categories.

– Creating a user page: It serves as the first step to belonging to the Wikipedia community and gets the users started with editing a Wikipedia page which includes personal information. It provides a practice experience without too much concern regarding the content of the page.
– Editing in Sandbox: Wikipedia provides Sandbox pages as a practice environment for users to practice with syntax of the Wiki Markup language to edit Wikipedia pages as well as organizing the content of the page before editing the main article page.
– Editing article pages: It indicates a stronger level of readiness for editing in Wikipedia and a stronger involvement in Wikipedians community

To study our second research question, we identified interactions happening on Wikipedia talkpages among edit-a-thons newcomers. Following the approach in [16], we excluded talkpage posts made by automatic Wikipedia bots. We constructed a communication network based on the talkpage interactions. In addition to interactions happening on Wikipedia, participants are encouraged to post about the event and communicate on Twitter using #ArtAndFeminism and #NowEditingAF hashtags. We utilized the Twitter interactions as representation of the social interactions as well. To study our third research question, we conducted a regression analysis to predict the relationship between various factors of the edit-a-thon events and subsequent participation of newcomers in Wikipedia.

5.1 Wikipedia dataset

In addition to the Wikipedia data related to the face-to-face edit-a-thons, we constructed two additional dataset as analogous group of newcomers to compare against edit-a-thons newcomers participants. Below, we provide information about each data collection.

– **Face-to-face Art+Feminism edit-a-thons**: This dataset included data from 59 edit-a-thons event happening from Jan 1, 2016 until March 5, 2016. Each edit-a-thon has a dedicated Wikipedia page associated to the event which includes the list of the participants [9]. We collected the list of participants from the Wikipedia pages. Using the Wikipedia API, for each participants of the edit-a-thons, we collected the

---

[9] For example: `https://en.wikipedia.org/wiki/Wikipedia:Meetup/Florida/ArtAndFeminism_2016`

day they had registered on Wikipedia, and all of their Wikipedia edits until April 2, 2016 (last day of our data collection).

– **Virtual Art+Feminism edit-a-thons**: In addition to face-to-face events, Wikipedians interested in improving representation of Feminism related articles and female editors, have been organizing virtual edit-a-thons[10]. Anyone from anywhere in the world can participate in the virtual events. The virtual edit-a-thon events were organized over period of two weeks or a month. We collected data from four virtual edit-a-thons happening around the same time as our face-to-face edit-a-thon collection from Dec 2015 until March 2016. Those interested in participation were encouraged to sign up online on the hosting page which included information about the facilitators and the list of articles to work on.

– **Randomly selected newcomers**: Our last collection of newcomers was a set of randomly selected newcomers on Wikipedia. For each newcomers attending a face-to-face edit-a-thon event we randomly selected a group of 10 newcomers who joined Wikipedia on the same day, then for each newcomer in this dataset, we collected all their editing activities on Wikipedia.

5.2 Twitter dataset

Finally, we collected the interactions happening on Twitter related to the ArtandFeminism Edit-A-Thons. For this purpose we collected all tweets from 2016 that contained the hashtags of #ArtAndFeminism and #NowEditingAF; or any of the hashtags reported on Wikipedia to be representative of a related edit-a-thon. This dataset included a total of 3,341 tweets from 1,171 different users related to 59 edit-a-thons events happening from Jan 1, 2016 until March 5, 2016. We also gathered information about other twitter users mentioned in the tweet, the text of the tweet, the date in which the tweet was posted, and any additional hashtag associated with those tweets. We also included the general Twitter interactions of edit-a-thon participants before, during, and after the event, other than their Art+Feminisim related tweets. We collected all the tweets they had posted for a period of 15 days before and after the event. We collected these tweets to be able to assess the level of connections between the participants on Twitter before and after the events by identifying user mentions in the tweets.

6 Results

6.1 RQ1: impact of Edit-A-Thons on subsequnet participation of newcomers

Among the edit-a-thon participants, we defined users as newcomers if they had not edited Wikipedia before the edit-a-thon event. The dataset included total of 1,018 participation from 985 unique participants with 586 (60%) of them identified as newcomers. The number of participants per each event ranged from 3 to 131 with average of 17.25 (Std. Dev = 19.47). The proportion of newcomers in each event ranged from 0 to 100% with average of 57% (std. Dev. 23.5%). All together, a total of 793 articles were

---

[10] `https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Women_in_Red/`
`Meetup/8`

edited during these edit-a-thons events with 475 (60%) articles edited by newcomers. There were total of 2,928 edits made to these 793 articles with 1,579 (54%) of them made by newcomers. 119 out of 793 articles were edited at least once after the events by one of the participants. Overall, an important observation of the data is the large percentage of newcomers attending each of the events. In fact, as mentioned earlier, we also collected data on virtual edit-a-thons for Art+Feminism. The dataset includes a total of 182 participation from 118 unique users. All except two of these users had been already registered on Wikipedia prior to the virtual edit-a-thon events and only 4 who had less than one edit prior to the events that we could consider as newcomers. These results suggest that the face-to-face events are much more likely to attract newcomers than the online events.

Our data shows that among the group of newcomers joining Wikipedia on the same time period as participants of the edit-a-thons, only 1% of the newcomers edit Wikipedia a week after registering on Wikipedia. To confirm our results, we repeated this analysis with three different randomly selected group of newcomers. In each case, we randomly selected a group of 5,233 newcomers and among them between 48 (0.9%) to 53 (1%) newcomers had edited any Wikipedia pages at least one week after registering on Wikipedia. On the other hand, among the 586 newcomers attending our target Edit-A-Thons event, 50 (9%) of them continued editing Wikipedia pages a week after the edit-a-thon event. Our results suggest that, compared to randomly selected newcomers on Wikipedia, a significant larger percentage of edit-a-thons participants continue editing Wikipedia pages. However, we would also like to acknowledge that randomly selected newcomers provides a baseline benchmark for comparison but it does not provide a fair comparison in terms of motivational factors and identification with the topic of Wikipedia articles. It is very likely that edit-a-thons participants have a stronger identification with topics represented during the edit-a-thon events that can encourage their subsequent participation as well.

Table 1 shows descriptive statistics on newcomers' activities on the day of the edit-a-thon event. The results shows that editing article pages was the most common activity and that a large number of newcomers did not get involve in editing any other Wikipedia pages, including creating their user pages.

Table 1. Newcomers activities on Edit-A-Thon event day

	% of users	Average	Std. Dev.	Median
User page	33%			
Article edits	48%	2.45	4.36	1
Sandbox edits	21%	.86	3.56	0
Other edits	11%	2.54	3.49	2

Summary of results In summary, in response to our first research question on impact of edit-a-thons on newcomers' participation, our results shows that overall face-to-face

edit-a-thons are very successful in attracting and recruiting a large number of newcomers who are more engaged than a random group of newcomers on Wikipedia; however, still a very small percentage of them stay engaged with Wikipedia after the event. Given somewhat limited activity level of newcomers on the event day, one potential solution to achieve more sustained engagement can involve encouraging newcomers to get involved on various Wikipedia editing activities during the event, especially activities such creating user pages that is more an entry level activity while connecting newcomers to the community.

6.2 RQ2: impact of Edit-A-Thons on newcomers' community connections

To further our understanding of the impact of Edit-A-Thons, beyond production mechanisms, we were interested to investigate the social interactions of the participants during and after the events. As mentioned earlier, we employed talkpage interactions and Twitter interactions to construct the interaction networks of newcomers.

Figure 1 presents newcomers' outgoing talkpage communication network; i.e. newcomers and all those individuals with whom the newcomers communicated on their talkpages and figure 2 presents newcomers' incoming talkpage communication network; i.e. communication they received from others on their talkpages. the network in Figure 1 is generated by extracting all the posts made to talkpages by any of the newcomers. The target could be another newcomer or an existing member. The network in Figure 2 is generated by extracting all the posts made to newcomers' talkpages and the source can be either a newcomer or an existing member. In each network, the light green nodes indicate the newcomers and the dark green nodes represent the existing members of Wikipedia. The thickness of the edges represent the number of talkpage post by that user. As presented in Figure 1, very few newcomers post on others' talkpages. Furthermore, as presented in Figure 2, communication between existing members and newcomers is also very limited and majority of newcomers have received very few messages on their talkpages and from a very few number of existing members. Since talkpages are the major place for communication and coordination among Wikipedia editors, this results suggest very little followup and engagement strategies employed by the existing Wikipedia members to keep these newcomers population engaged. The incoming network includes 665 nodes with 2.1 average number of neighbors and network density of .003. The outgoing network includes 64 nodes with 1 as average number of neighbors and network density of .02.

Next, We used the Twitter data to further unravel how edit-a-thon participants connected with each other. We constructed their connection network by defining connection between two user when one of them mention the other user explicitly in their Tweets associated with either #ArtAndFeminism and #NowEditingAF hashtags. Given that people can connect with many different users on Twitter, we limited the tweets to these two specific hashtags to detect the social interactions related to the edit-a-thons. We considered a mention as an interaction because a mention is an explicit indication that the two users have exchanged tweets and have interacted with each other. We also considered two users to be connected together as long as they had both been tweeting with the Art+Feminism hashtag and mentioned each other even if the mention was in tweets
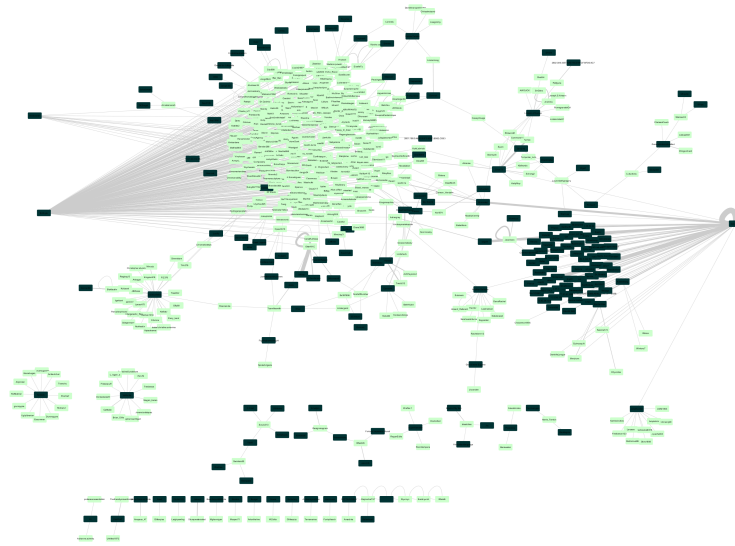
Fig. 1. Newcomers' outgoing communication network

not including those specific hashtags. Figure 3.b presents a visualization of the connections between participants of one the edit-a-thons with the most online interactions, the March 5th,2016 edit-a-thons.

Figure 3.b shows that a small group of people interacted with each other even during the event day, forming hubs. We observe that a fair number of isolate nodes appeared. These are individuals who were tweeting a lot during the event without connecting to others. We also observed that a very few accounts were mentioned for a very large number of times. Upon further manual inspection, we found that these accounts belonged to established organizations usually running the edit-a-thons or providing a space for the edit-a-thon events (e.g., the account muac_unam was one of the most mentioned in edit-a-thons. This account belongs to a museum at the National Autonomous University of Mexico (UNAM), and has organized some of the biggest edit-a-thon events in Latin America.) This finding hints that the usage of Twitter during the edit-a-thons have been focused on more of official communication than bonding and connection building among participants. There seems to be little attempt to engage the general participants and newcomers of the edit-a-thons through twitter messages. An example of this is the following tweet from @Wikimedia_mx mentioning UNAM museum during an Edit-A-Thon:

"We continue editing about art and women at @muac_unam #ArtAndfeminism"

The following example is from a user tweeting about participating in an edit-a-thon at the Menil Collection Library:

"Art + Feminism wiki Edit-A-Thon. #NowEditingAF in the @menilcollection library"

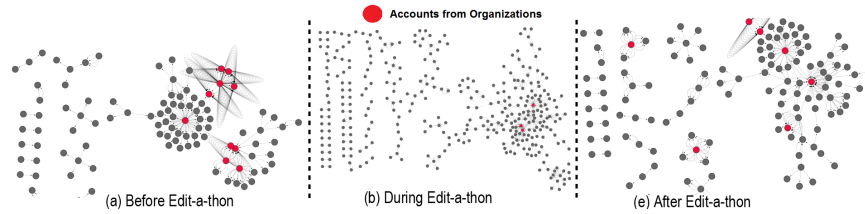**Fig. 2.** Newcomers' incoming communication networks

Fig. 3. Twitter network for March 5,2016 Edit-A-Thon

Table 2 present the descriptive statistics about the general Twitter activity of our target participants. We observe that participants were fairly active on Twitter, connecting also with other individuals; however, their presence on Twitter seems to have not been utilized in relation to edit-a-thon events and connecting with other participants.

Figures 3.a and 3.c present the Twitter communication network before and after the edit-a-thon events. The network is generated by considering all Twitter interactions of the participants before and after the event. The after-network highlights that only the main organizations (highlighted in red) involved in the edit-a-thons are the ones that people most reach out to; the general participants seem to be lost from the communication network. This further confirms our observations that although the events are able to attract a large number of newcomers, the interaction among them is very low, and communication stays within these more influential and well established accounts that represent organizations.

We also analyzed the type of hashtags that edit-a-thon participants used before, during, and after the event. Perhaps, although the participant were not reaching out to

Table 2. Summary of general Twitter activity of participants before, during, and after the Edit-a-thons events

| | Before | | During | | After | |
	Median	Average	Median	Average	Median	Average
Tweets	11	26	2	3	9	26
Hashtags	12	36	4	9	14	42
Mentions	22	68	3	9	21	67

each other, they might keep a certain bond and connection to the group by tweeting using edit-a-thon related hashtags. Table 3 presents some of the most popular hashtags used before, during, or after the edit-a-thons. One of the most popular hashtags was #ArtAndFeminism. However, we observed that the hashtag was used primarily before and during the edit-a-thon and not after the event which confirms our prior observation of lack of follow through after the event.

On the other hand, we observed that some of the most popular hashtags used by involved users in the tweets not associated with the official hashtags (before, during, and after the event) still appeared to be related to Feminism (e.g., *#MyFeminismIs*, *#FeministFriday*, *InternetFeminista*, *#SadFeministCat*). However, those hashtags were not adopted by the edit-a-thon organizers that could have been utilized to further engage and motivate participation of highly-motivated individuals in Wikipedia. As shown in Table 2, in general the participants are rather active on Twitter but not very engaged with the specific edit-a-thon hastags. All together, our results suggests that Twitter might be a useful platform to be employed for engaging edit-a-thon participants after the event.

Table 3. Associated hashtags used before, during, and after by Edit-A-Thons involved Twitter users

	Hashtags
Before	#editatón, #editathon, #5womenartists, #Wikipedia, #8demarzo, #gendergap, #feminist, #artandfeminism, #WomenInRed
During	#editatón, #editathon, #5womenartists, #Wikipedia, #8demarzo, #gendergap, #feminist, #artandfeminism, #WomenInRed
After	#microaggressions, #feministsplaining, #homosinherstory, #artlibrariessowhite, #archivistproblems

Summary of results In summary, in response to our second research question on impact of edit-a-thon on forming connections, we observed very little evidence on that on either Wikipedia talkpages or Twitter. We observed very little social connections created among the participants and very little followup after the events to further engage the newcomers. At the same time, we observed that many of the participants of edit-a-thon events are active Twitter users and exhibited particular interest on the topic of

feminism. These results together hints at possibly a missed opportunity that can be utilized by Wikipedia community and organizers of edit-a-thons to increase the liklihood of newcomers' engagement after those events.

6.3 RQ3: predicting subsequent participation

While overall we observed low retention rate among edit-a-thon participants, we were interested to assess whether any of the production and social mechanisms were related to higher likelihood of retention and subsequent participation. To do so, we conducted a repeated measure logistic regression analysis to predict whether the newcomers would edit any Wikipedia pages at least a week after the event. The model nested individual users within the edit-a-thon event they attended. We included the production measures of creating a user page, editing article pages, or sandbox pages, the number of participants in the event, as well as the proportion of newcomers attending the event as independent variables in the model. In terms of interaction mechanisms, we utilized the talkpage communication network and for each newcomers, we calculated common network measures of closeness centrality, betweenness centrality, degree, and clustering coefficient. Degree represents the number of immediate connections a node has in the network and indicates how well-connected a node is; Closeness centrality represents a more global level of connectedness in the network through considering the distance of a node to all others in the network. Betweenness centrality on the other hand focuses on favored bridging positions of nodes and how many paths of connections in the network rely on this particular node. Clustering coefficient represent how close a node's neighbors are to being completely connected graph. However, in our dataset the measure of degree was significantly correlated with all other network measures (Table 4; therefore, to avoid multicollinearity, we only included degree in the regression model. Since the size of the network and degree can be related, we ensured that degree was not correlated to the number of participants (coef=.02, Sig.=.71).

Table 4. Correlation between network measures

	betweenness		closeness		clustering	
	coef	Sig.	coef	Sig.	coef	Sig.
degree	.84	<.001	.23	<.001	.29	<.001

The result of the regression analysis for the significant factors is presented in Table 5. The results show that number of participants, editing articles, and receiving talkpage messages are correlated with higher likelihood of continuing to edit Wikipedia. Any one additional person attending the event leads to 1% increase in the odds of having a newcomer continue editing a Wikipedia page. An additional talkpage message leads to 68% increase in the odds of having a newcomer continue editing a Wikipedia page and one additional edit in article pages on the event day increases the odds of having a newcomer continue editing a Wikipedia page by 7%. The effect of editing articles

is only marginally significant. Overall the results suggest that on-event support and social interactions can lead to higher level of subsequent commitment. Other factors, including proportion of newcomers, editing sandbox pages, and creating user pages were not significant factors in the model.

Table 5. Analysis result

	Odds ratio	Std. Error	Sig.
Number of participants	1.01	.004	.021
Degree	1.68	.16	.001
Article edits	1.07	.03	.079

7   Conclusion and discussion

In this work, we presented a case study of the impact short-term collocated event, that are quite popular among various organizational groups and online communities, affect on onboarding of newcomers into the community. We studied this in the context of Art+Feminism Wikipedia edit-a-thons focusing on increasing representation of female editors. To understand the production and social mechanism of these edit-a-thons, we triangulated different sources of log data on Wikiepdia and Twitter. Our results show that these events are very successful in attracting new members. A significant percentage of participants in each of those events are individual with no prior Wikipedia experiences, at a much higher rate in comparison to their parallel virtual events ; however, they are not very successful in retention of these motivated individuals. We speculate that higher level of hands-on activities on the event day, and followup communications and engagement mechanisms can play significant role in increasing retention.

Our analysis of newcomers talkpage communication network and participants' Twitter communication speak to these speculation. We found very little communication happening on Wikipedia and we found that other than official organizational accounts associated with the edit-a-thon events, very little followup communication happened on Twitter. On the other hand, similar to previous research we observed the importance of interaction between existing members and newcomers on encouraging their future participation. Receiving messages on talkpages was associated with a significant increase in odds of future contribution of a newcomer. Moreover, we observed that when the general participants tweeted they use a wider range of hashtags which were somewhat disjoint from hashtags utilized by administration members. This can signify that potential strong motivation of participants which might have not been capitalized by those in charge of these events.

At the same time, our results also shows that activities on the day of the event can be important in encouraging future participation. We did not observe any support for impact of creating user pages or practicing editing in Sandbox pages on future participation; however, those actions were not common among the participants and there were

very small occurrence of those cases that could affect the result of our analysis. Encouraging to take on such actions which can be a good starting point can be beneficial in encouraging further participation and fostering sense of belonging to the community.

Our analysis are based on archival log data and in the future we plan to conduct interviews with organizers and participants to gain a deeper insight into results highlighted in our current work and better understanding of the goals of each event and their satisfaction with the extent they have achieved their goals. It is possible that in some of these events, on-event activities were of higher importance to the organizers than future participation of a large number of participants. Our initial contact with a few of organizers has been received with high enthusiasm, especially with regards to the issue of engagement of newcomers that they acknowledges as a challenge. Informed by the results of our current work, in the future, we plan to work closely with organizers of these events in experimenting various followup strategies to increase newcomers' retention into the online production community.

Nevertheless, our current work highlights the value of these offline gatherings on attracting a new stream of newcomers while providing insight on the challenges they face and potential approaches in addressing such challenges.

References

1. Ahuja, M.K., Galvin, J.E.: Socialization in virtual groups. Journal of Management 29(2), 161–185 (2003)
2. Angelopoulos, S., Merali, Y.: Bridging the divide between virtual and embodied spaces: Exploring the effect of offline interactions on the sociability of participants of topic-specific online communities. In: International Conference on System Sciences. pp. 1994–2002. IEEE (2015)
3. Brady, E., Morris, M.R., Bigham, J.P.: Gauging receptiveness to social microvolunteering. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 1055–1064. ACM (2015)
4. Burke, M., Marlow, C., Lento, T.: Feed me: motivating newcomer contribution in social network sites. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 945–954. ACM (2009)
5. Chen, J., Ren, Y., Riedl, J.: The effects of diversity on group productivity and member withdrawal in online volunteer groups. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 821–830. ACM (2010)
6. Choi, B., Alexander, K., Kraut, R.E., Levine, J.M.: Socialization tactics in wikipedia and their effects. In: Proceedings of the 2010 ACM conference on Computer supported cooperative work. pp. 107–116. ACM (2010)
7. Ciampaglia, G.L., Taraborelli, D.: Moodbar: Increasing new user retention in wikipedia through lightweight socialization. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. pp. 734–742. ACM (2015)
8. Cohen, N.: Define gender gap? look up wikipedia's contributor list. New York Times, January 30(362), 1050–56 (2011)
9. Ducheneaut, N.: Socialization in an open source software community: A socio-technical analysis. Computer Supported Cooperative Work (CSCW) 14(4), 323–368 (2005)
10. Farzan, R., Kraut, R.E.: Wikipedia classroom experiment: bidirectional benefits of students' engagement in online production communities. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 783–792. ACM (2013)

11. Faulkner, R., Walling, S., Pinchuk, M.: Etiquette in wikipedia: Weening new editors into productive ones. In: Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration. p. 5. ACM (2012)
12. Ganglbauer, E., Fitzpatrick, G., Subasi, Ö., Güldenpfennig, F.: Think globally, act locally: a case study of a free food sharing community and social networking. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. pp. 911–921. ACM (2014)
13. Halfaker, A., Geiger, R.S., Morgan, J.T., Riedl, J.: The rise and decline of an open collaboration system: How wikipedia's reaction to popularity is causing its decline. American Behavioral Scientist p. 0002764212469365 (2012)
14. Halfaker, A., Kittur, A., Riedl, J.: 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. pp. 163–172. ACM (2011)
15. Jin, L., Robey, D., Boudreau, M.C.: The nature of hybrid community: An exploratory study of open source software user groups. The Journal of Community Informatics 11(1) (2015)
16. Johnson, I., Lin, Y., Jia-Jun Li, T., Hall, A., Halfaker, A., Schöning, J., Hecht, B.: Not at home on the range: Peer production and the urban/rural divide (2016)
17. K. Lam, S., Uduwage, A., Dong, Z., Sen, S., Musicant, D.and Terveen, L., Riedl, J.: Wp:clubhouse?: an exploration of wikipedia's gender imbalance. In: WikiSym'11. ACM (2011)
18. Kraut, R., Burke, M., Riedl, J., Resnick, P.: Dealing with newcomers. Evidence-based Social Design Mining the Social Sciences to Build Online Communities 1, 42 (2010)
19. Lam, S.T.K., Uduwage, A., Dong, Z., Sen, S., Musicant, D.R., Terveen, L., Riedl, J.: Wp: clubhouse?: an exploration of wikipedia's gender imbalance. In: Proceedings of the 7th International Symposium on Wikis and Open Collaboration. pp. 1–10. ACM (2011)
20. Lampe, C., Johnston, E.: Follow the (slash) dot: effects of feedback on new members in an online community. In: Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work. pp. 11–20. ACM (2005)
21. Lin, H.F.: The role of online and offline features in sustaining virtual communities: an empirical study. Internet Research 17(2), 119–138 (2007)
22. Morgan, J.T., Bouterse, S., Walls, H., Stierch, S.: Tea and sympathy: crafting positive new user experiences on wikipedia. In: Proceedings of the 2013 conference on Computer supported cooperative work. pp. 839–848. ACM (2013)
23. Panciera, K., Halfaker, A., Terveen, L.: Wikipedians are born, not made: a study of power editors on wikipedia. In: Proceedings of the ACM 2009 international conference on Supporting group work. pp. 51–60. ACM (2009)
24. Savage, S., Monroy-Hernandez, A., Hollerer, T.: Botivist: Calling volunteers to action using online bots. arXiv preprint arXiv:1509.06026 (2015)
25. Sessions, L.F.: How offline gatherings affect online communities: When virtual community members 'meetup'. Information, Communication & Society 13(3), 375–395 (2010)
26. Shaw, A., Zhang, H., Monroy-Hernández, A., Munson, S., Hill, B.M., Gerber, E., Kinnaird, P., Minder, P.: Computer supported collective action. interactions 21(2), 74–77 (2014)
27. Shen, C., Cage, C.: Exodus to the real world? assessing the impact of offline meetups on community participation and social capital. New Media & Society p. 1461444813504275 (2013)
28. Suh, B., Convertino, G., Chi, E.H., Pirolli, P.: The singularity is not near: slowing growth of wikipedia. In: Proceedings of the 5th International Symposium on Wikis and Open Collaboration. p. 8. ACM (2009)
29. Tewksbury, D.: Online-offline knowledge sharing in the occupy movement: Howtooccupy. org and discursive communities of practice. American Communication Journal 15(1), 11–23 (2013)

30. Trainer, E.H., Chaihirunkarn, C., Kalyanasundaram, A., Herbsleb, J.D.: Community code engagements: summer of code & hackathons for community building in scientific software. In: Proceedings of the 18th International Conference on Supporting Group Work. pp. 111–121. ACM (2014)
31. Trainer, E.H., Kalyanasundaram, A., Chaihirunkarn, C., Herbsleb, J.D.: How to hackathon: Socio-technical tradeoffs in brief, intensive collocation (2015)
32. Zhu, H., Zhang, A., He, J., Kraut, R.E., Kittur, A.: Effects of peer feedback on contribution: a field experiment in wikipedia. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2253–2262. ACM (2013)