Participation of New Editors After Times of Shock on Wikipedia

Ark Fangzhou Zhang
University of Michigan
arkzhang@umich.edu

Ruihan Wang
University of Michigan
ruihwang@umich.edu

Eric Blohm
University of Michigan
ericblo@umich.edu

Ceren Budak
University of Michigan
cbudak@umich.edu

Lionel P. Robert Jr.
University of Michigan
lprobert@umich.edu

Daniel M. Romero
University of Michigan
drom@umich.edu

Abstract

User participation is vital to the success of collaborative crowdsourcing platforms such as Wikipedia. Previously user participation has been studied during "normal times". However, less is known about participation following shocks that draw attention to an article. Such events can be recruiting opportunities due to increased attention; but can also pose a threat to the quality and control of the article and drive away newcomers. We study the collaborative dynamics of Wikipedia articles after times corresponding to shocks generated by drastic increases in attention as indicated by data from Google trends. We find that participation following such events is indeed different from participation during normal times–both newcomers and incumbents participate at higher rates during shocks. We also identify collaboration dynamics that mediate the effects of shocks on continued participation after the shock. The impact of shocks on participation is mediated by the amount of negative feedback given to newcomers in the form of reverted edits and the amount of coordination editors engage in through edits of the article's talk page.

Introduction

Sustained user participation is vital to the survival and success of Wikipedia articles, which rely on collaborative inputs from crowds of volunteers (Butler 2001; Halfaker et al. 2013). The participation of editors to an article has been identified as a major determinant of article quality (Ransbotham and Kane 2011; Robert and Romero 2017). While some articles naturally attract enough participant to contribute new content, provide copy editing, and ensure adherence to Wikipedia guidelines, others fail to attract enough volunteers and struggle to achieve high quality (Halfaker et al. 2013). Indeed, the distribution of the number of editors among Wikipedia articles follows a heavy tail distribution, where a small percentage of articles have a large number of contributors while most articles are edited by just a few editors (Voß 2005). Thus, increasing participation of editors in articles across the platform is an important problem to solve for the sustainability of Wikipedia.

One important mechanism that drives participation to an article is an *attention shock*—a real-world event that generates attention, in the form of readership and contributions, to a specific Wikipedia article (e.g. the death of a celebrity). These shocks can present an opportunity to an article as they attract large numbers of editors who can potentially become long-term contributors to the article. However, the influx of a large number of new editors can be overwhelming, possibly leading to conflict between editors (Halfaker et al. 2013; Keegan, Gergle, and Contractor 2013). The collaboration and communication dynamics during the shock might affect newcomers' perception of self-worth and in turn dictate their willingness to become long term contributors to the article.

We aim to understand the effects of attention shocks on the participation of new editors to an article. Importantly, we uncover the collaborative dynamics within an article that lead newcomers' continued participation after the shock. Much of the work on participation of newcomers in Wikipedia has focused on platform newcomers—users who are new to Wikipedia, and on the retention of such newcomers (Halfaker, Kittur, and Riedl 2011; Halfaker et al. 2013; Faulkner, Walling, and Pinchuk 2012; Mesgari et al. 2015; Suh et al. 2009; Li and Farzan 2018; Robert and Romero 2015; 2017; Chen, Ren, and Riedl 2010; Ransbotham and Kane 2011). In contrast, our study considers newcomers to an article who may have some experience editing other articles in Wikipedia. However, as we will show, articles newcomers during shocks have significantly lower Wikipedia experience than incumbents, suggesting that some of the same challenges in retaining platform newcomers (Halfaker, Kittur, and Riedl 2011) will apply to article newcomers as well. Our study also goes beyond measuring *whether* an article newcomer returns to the article (retention) and measure their participation by *how much* they contribute relative to the contributions of existing editors.

Given that most users contribute a small number of edits on Wikipedia, it is not enough to find ways for users to be active but also to contribute a significant amount of work.

We posit that attention shocks can have a direct and indirect effect on the participation of newcomers. The indirect effect can occur through changes in collaborative dynamics of the article during the shock. We analyze three types of collaborative features that are associated with levels of participation of members of online communities (Zhu et al. 2013a; Halfaker et al. 2013). Specifically, we focus on 1) negative feedback measured by reverts–revisions that undo other editors' contributions (Zhu et al. 2013a; Halfaker et al. 2013), 2) centralization of contributions as measured by distribution of editor efforts (Zhang et al. 2017a; 2017b), and 3) levels of coordination measured by discussions on talk pages (Romero, Huttenlocher, and Kleinberg 2015). Through mediation analysis, we quantify the direct and indirect effects of shocks and identify specific collaborative features that lead to higher newcomer participation.

To capture attention shocks that are exogenous to Wikipedia, we use Google Trends to identify Wikipedia entries that exhibit spikes in attention as reflected by related Google queries. In order to facilitate validation and increase the chances of finding a genuine attention shock, we focus on people who have enough notability to have an associated Wikipedia article. Within these articles, we conduct a separate analysis on academics, politicians, and a random sample of people who are listed under WikiProject Biography. Our results are consistent across the three sets of articles.

Our analysis shows that shocks on Wikipedia articles tend to bring a large number of new editors to articles. Additionally, attention shocks have a consistent, positive but temporary impact on participation of newcomers after the shock. This impact is partially mediated by reverts of newcomers and edits on talkpages. In sum, we find that shocks facilitate participation in Wikipedia crowds by promoting reverts of newcomers together with increasing edits on talkpages. Furthermore, there is a significant spillover effect on the retention of newcomers over all of Wikipedia. In other words, the newcomers who join an article during times of shock are more likely to contribute more *to other Wikipedia articles* compared to newcomers that join during normal times.

Our study sheds light on the dynamics of collaboration on Wikipedia among new and existing crowd members during periods of large influx of editors. Understanding these dynamics is crucial to effectively recruit participants to online communities and crowd systems such as Wikipedia articles when they are driven to participate by exogenous events.

Related Work

Measures of Participation in Wikipedia Participation is the degree to which members contribute to the creation and maintenance of Wikipedia pages. This contribution can come in many forms and quantities. Examples include contributing to talk pages, reverts or adding new text. Prior studies have measured participation is various ways such as tracking whether participants continue contributing after they initially join (i.e. retention or continued participation) (Karumur et al. 2018; Yu et al. 2017) and measuring how much they contribute (Chen, Ren, and Riedl 2010; Zhu et al. 2013b). In Wikipedia studies, participation has often been used to explain crowd success. More specifically, prior research has found a positive association between various measures of member participation and article quality—highlighting the importance of this measure (Arazy and Nov 2010; Jones 2008; Yu et al. 2017). In this study, we measure participation in several ways, but always consider the amount of participation by the users rather than simply using a binary measure of whether they contributed at least one edit.

Platform and Crowd Participation. An important question in crowdsourcing systems such as Wikipedia is how to encourage continued participation of users. However, the definition of "group" can be defined by membership or participation in an article, in a WikiProject, or in Wikipedia as a whole (Platt and Romero 2018). The literature on retention can be divided into two areas: platform and crowd retention (i.e. continued participation). Platform retention focuses on retaining new members to the Wikipedia platform. These studies focus on various approaches to attracting and keeping new platform members (Halfaker et al. 2013; Morgan and Halfaker 2018; Schneider, Gelley, and Halfaker 2014; Li and Farzan 2018). An example is Halfaker et al. (2011), which examined the impact of platform policies and norms on the retention of new members. Perhaps more closely related to our study, Li and Farzan (2018), studied the behavior of editors that join Wikipedia during three *current events*. Crowd retention focuses on retaining new members to a particular Wikipedia crowd, such as in individual article or Wikiproject within the platform. These studies acknowledge that current platform members can and do choose to contribute to certain articles but never contribute to others (Chen, Ren, and Riedl 2010; Karumur et al. 2018; Ransbotham and Kane 2011; Zhu et al. 2013b).

These studies examine the impact of crowd retention on individual and group participation and/or some measure of article quality (Chen, Ren, and Riedl 2010; Karumur et al. 2018; Ransbotham and Kane 2011; Zhu et al. 2013b). For instance, Robert and Romero (2017) examined the impact of crowd retention on article quality relative to crowd size. On occasion, scholars have considered both platform and crowd retention (Yu et al. 2017).

Our study seeks to add to the conversation on *crowd participation* by exploring the impact of a large number of attention shocks on the participation of edits who are new to a Wikipedia article regardless of their prior experience on the Wikipedia platform.

Shocks. Current research on participation, at the crowd or platform level, has focused on "normal" or non-shock times. Unfortunately, we know little about the impact of shocks on crowd future participation. This is despite the fact that crowds operate in a rapidly changing environments often impacted by external events with little forewarning (Zhang et al. 2017a). The lack of formal boundaries and the ability of crowd members to leave at any time means that crowd participation are likely to be particularly susceptible to shocks (Robert and Romero 2017). Building on prior work, we address these questions and provide new insights to the literature on crowd participation.

Example of prior research that study Wikipedia dynamics in the context of shocks are (Keegan, Gergle, and Contractor 2012; 2013), which analyze how articles and editors respond to breaking news events and (Leung, Zhu, and Konstan 2017), which studies whether National Football League game outcomes affect the editing behavior of Wikipedia editors who identify as fans of a team. However, the shocks studied in these papers are typically not associated with an existing Wikipedia article. In contrast, the events we study are mostly related to figures who are already notable enough to have a Wikipedia article before the shock. Thus, shocks will not only bring in newcomers, but also impact the behavior of incumbent editors whose behavior can influence newcomers.

Hypotheses Development

Here, we construct and summarize a theory-driven framework to form hypothesis as to how shocks affect collaboration dynamics, and how participation changes directly as a function of the shock as well as due to the changes in collaboration dynamics (Fig 1). We now describe our hypotheses.

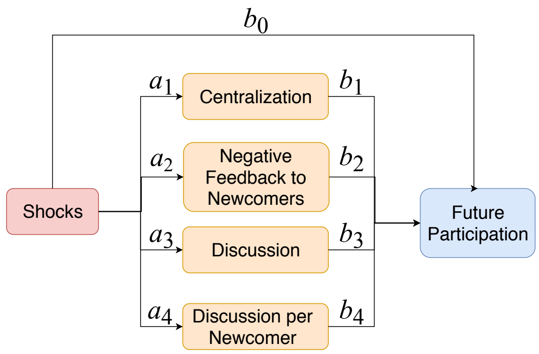


Figure 1: Mediation Model

Attention Shocks and Collaboration Dynamics

Shocks and Centralization. Theory suggests competing hypotheses for the effect of attention shocks on centralization — the extent to which an article is controlled and produced by a small fraction of editors. Threat rigidity would suggest that crowds would react to attention shocks by becoming more centralized (Staw, Sandelands, and Dutton 1981; Zhang et al. 2017a). Yet, an influx of new article editors unfamiliar with work practices and norms would make it difficult to maintain the current social order (Buechler 2013; Useem 1998), which involves maintaining a highly centralized structure (Romero, Huttenlocher, and Kleinberg 2015; Kittur and Kraut 2008). Thus, we expect a relationship between shocks and centralization, but we do not make a hypothesis on the direction of the effect.
H1: Shocks are related to centralization.
Shocks and Negative Feedback to Newcomers. Attention shocks could have a significant impact on the rate of reverts of article newcomers. As well will show, article newcomers during shocks tend to have far less experience editing Wikipedia than incumbents, making them less aware of Wikipedia level norms and policies. Moreover, since they are new to the article, they will not be familiar with article specific norms. Thus, newcomers to articles are less likely to understand crowd norms (Krieger, Stark, and Klemmer 2009), and existing members have less time to train them. Additionally, it is possible that existing editors will feel threatened by newcomers and express territoriality by rejecting their contributions.
H2: Shocks are positively related to reverts of newcomers.
Shocks and Discussion. Attention shocks bring more work and more newcomers to an article. This should increase the need for coordination among editors.
H3: Shocks are positively related to edits to talkpages.
Shocks and Newcomer's Contributions to Discussions. Increases in coordination during shocks mentioned above will also lead to more discussion by newcomers.
H4: Shocks are positively related to newcomer's edits to talkpages.

Collaboration Dynamics and Participation

Centralization and Participation. Centralized Crowds present a less chaotic and more ordered work environment (Romero, Huttenlocher, and Kleinberg 2015). For article newcomers, a highly ordered environment is easier to navigate and presents a more attractive place to stay and continue to participate (Choi et al. 2010a; Robert and Romero 2017).
H5: Centralization is positively related to participation.
Negative Feedback and Newcomer Participation. Reverts often represent conflict and disagreement, which might discourage newcomers from staying and participating (Robert and Romero 2017). Reverts of newcomers can also represent mentoring and feedback which may encourage newcomers to stay and participate (Zhu et al. 2013a). Thus, we have competing hypotheses regarding the relationship between reverts and participation.
H6: Reverts of newcomers are related to participation.
Discussion and Participation. Group discussion is often vital to socializing new editors. Edits to talk pages represents a vibrant discussion among crowd members. This socialization process represents an opportunity for newcomers to observe and learn crowd norms and work practices (Ciampaglia and Taraborelli 2015). Both are vital to encouraging participation (Choi et al. 2010b).
H7: Discussion is positively related to participation.
Newcomers' Revisions on TalkPages and Participation. Newcomers who edit talk pages are able to socialize with other editors and learn community norms. This will make them more likely to continue editing in the future.Therefore, we expect contributions to talk pages to lead to newcomers staying and participating more.
H8: Newcomer's contribution to discussion are positively related to participation.

Finally, based on the previous hypotheses, we expect that shocks will have an impact on participation through several collaboration features.
H9: The impact of shocks on participation is partially mediated by centralization, reverts of newcomers, edits to talkpages and newcomer's edits to talk pages.

Wikipedia Shocks

We first identify events that may trigger a rapid increase in attention to Wikipedia articles—focusing on Wikipedia articles about people since these articles are susceptible to changes when individuals are involved in events such as deaths, scandals, or electoral victories. Furthermore, we conduct analysis and compare the results for three types of individuals: (i) academics, (ii) politicians, and (iii) a random sample of articles from *WikiProject Biography*, which contains over 1.4 million articles about people.

Google Trends

In order to identify events that cause a rapid increase in the attention on the person associated with the article, we use Google Trends (GT), a tool that provides a time series of Google search volume for a query relative to the maximum search volume on a given time interval. We obtain the time series of the volume of search for each name in our dataset from January 1, 2004 to July 31, 2017. We expect that increased search volume will lead to traffic to the person's Wikipedia article and, in turn, to an increase in the number of people editing the article (Antin and Cheshire 2010).

Given a search query, GT provides suggestions on topics associated with the query. For queries corresponding to people, the topic suggestion is often the profession of the person. For example, for the query "John Snow" GT suggests "Physician" and "Cricketer", and provides separate search time series for each profession. We match GT time series corresponding to (name, topic) pairs to Wikipedia articles following two criteria: (i) the name of the person matches the title of the Wikipedia article, and (ii) the topic appears in the first paragraph of the Wikipedia article. Since the topic is usually the profession of the person, this topic is likely to appear in the general description on their Wikipedia article.

Our matching procedure aims to increase our precision by requiring that the topic is mentioned in the article's first paragraph. While this strict requirement potentially leads to a loss in genuine matches of articles to GT time series, it also dramatically decreases the rate of false matches in our final sample. Since our goal is to analyze a reliable set of articles that were genuinely exposed to attention shocks, we optimize for high precision at the expense of recall.

Identifying attention peaks

Given a time series GT_t^a for article a, we identify times t^* with unusually high $GT_{t^*}^a$ with a two-step procedure:

Step 1: To control for trends and seasonal patterns, we perform time series decomposition using Loess (STL) to extract trend, seasonal, and remainder components (Cleveland, Cleveland, and Terpenning 1990). STL is a standard approach that decomposes an observed time series into several components. Specifically, it assumes that a time series is given by

$$y_t = T_t + S_t + RGT_t,$$

where T_t and S_t represent time trend and seasonality, respectively. The time trend and seasonality are driven by non-transitory factors such as a general increase in public's interest over all articles in the topic, while the remainder components are due to unexpected shocks. By interpolating data point with Loess local-smoothing on nearby time trends, STL separately identifies time trend, seasonality and remainder component.

Step 2: Using the remainder component of the time series, RGT_t^a, we first fit the values of the time series to a normal distribution $N(\mu, \sigma)$ in order to smooth out the time series. We then identify times t^* such that $RGT_{t^*}^a > Q_3 + \beta IQR$, where IQR is the interquartile range of RGT_t^a, Q_3 is the third quartile of RGT_t^a, and β is a constant. In our analysis we choose $\beta = 5$. This value is selected by minimizing the rate of false positives (shocks that do not clearly correspond to news events) for values $\beta = 3, 4, 5$. For each value of β, we select a random sample of 50 shocks and use Google search to identify news that can validate the shock. The rate of unvalidated shocks is 68%, 58%, and 24% for values $\beta = 3, 4, 5$, respectively. Given that we are unable to validate the majority of shocks for $\beta = 3, 4$, and that the rate of false positive drops significantly for $\beta = 5$, we choose $\beta = 5$ for our analysis.

Times t^* represent the peak of attention towards article a.

Since we remove the trend and the seasonality of the time series, we expect that t^* correspond to attention shocks, which in most cases are isolated or unexpected [1]. Note that t^* corresponds to the time when attention peaked but not necessarily when the event occurred–some lag can be expected between the event and peak of attention.

Table 1 shows the number of articles from the three groups: academic, politicians, and random sample and how many had a GT topic, were matched to a GT time series, and had an identified shock. In total, we consider over 275,000 articles and identify 6,662 attention shocks.

Category	#art.	#art. w/ a GT topic	#art. matched	#art. w/ shocks
Academics	64,362	25,295	15,241	936
Politicians	126,559	55,677	27,915	3,138
Biography	85,000	57,030	22,179	2,588
Total	275,921	138,002	65,335	6,662

Table 1: Number of articles with GT topics, matched to a GT time series, and with identified shocks by category.

Validation and Characterization of Shocks

In order to assess the validity of shocks, we match spikes in attention from GT to notable events pertaining to the highlighted individual. We select a random sample of 100 shocks for each type of individual and perform a Google search for news articles referencing the individual's name, published in the month that the shock occurred. Shocks are then classified into three categories: **(i.) Validated:** Shocks for which a notable event involving the individual was identified. **(ii.) Mislabeled:** Shocks for which a notable event not involving the individual was identified – typically due to a shock

[1] GT limits the granularity of the time series as a function of the requested time range. The time series is aggregated at a weekly level for ranges greater than 9 months; and at a monthly level for rangers greater than 5 years. Since we are interested in a large time range and the identification of fine-grained shocks (at the level of a day), we make two types of GT queries. First, we obtain a GT time series for 1/1/2004-7/31/2017 aggregated at the monthly level and run the two-step shock identification procedure on this time series. This allows us to identify the month when the shock occurred. We then obtain a daily level GT time series for the month of shock and define the shock to be the day with the highest search volume.

related to an individual with the same name. **(iii.) Not Identified:** Shocks for which no notable event identified.

Table 2 shows the number of shocks in each of these categories broken down by type. Among all, we validate 76%, we do not find relevant news for 16%, and we confirm that only 8% were mislabeled. This suggests that the precision of our procedure to find attention shocks is at least 0.76. The Politician category has the highest precision with 88% of shocks verified and the Academics category has the lowest with 66% of shocks verified. This is reasonable since politicians are generally more likely to be covered in the news and on the Web than academics.

Category	Validated	Mislabeled	Not Identified
Academics	66	13	21
Politicians	88	5	7
Biography	73	7	20
Total	227(76%)	25(8%)	48(16%)

Table 2: Number of shocks validated by category, from a random sample.

We further categorize verified shocks by the nature of the events. We identify 7 general categories that contain 83% of all verified shocks : (i) individual dies (Death); (ii) individual is involved in a scandal (Controversy); (ii) individual participates in an election (Election); (iii) individual is featured in the media (Media Feature); (iv) individual participates in a sporting competition (Sporting Event); (v) individual reveals their scientific discoveries (Scientific Discovery); (vi) individual wins a prize or award (Award); (vii) Other.

Table 3 shows the number of shocks validated in the 8 categories. News of an individual's death precedes a shock for all groups, while elections and controversies cause a shock for politicians and discoveries and awards increase attention towards academics. Overall, the attention shocks we find are triggered by a diverse set of events. This suggests that the results of our analysis generalizes to a wide variety of exogenous events.

Category	Academ.	Poli.	Bio.	Total
Death	19	17	11	47
Controversy	5	22	10	37
Election	3	31	2	36
Media Feature	9	1	18	28
Sporting Event	-	-	15	15
Scientific Discovery	12	-	1	13
Award	9	-	3	12
Other	9	17	13	39
Total	66	88	73	227

Table 3: Number of identified shocks following different types of events, across WikiProject Biography categories.

Finally, one may be concerned that our methodology finds shocks that are not exogenous, but endogenously driven by expected events like the planned released of a new product. In order to measure the exogeneity, we follow the approach from (Crane and Sornette 2008) and compute the parameter p of each shock. There are three classes on shocks that correspond to different values of p: exogenous subcritical ($p \approx 1.4$), exogenous critical ($p \approx 0.6$), and endogenous critical ($p \approx 0.2$). We find that 97.3% of our shocks have $p > 0.2$ and 83.4% have $p > 0.6$. Thus, we conclude that the vast majority of our shocks would be classified as either exogenous subcritical or exogenous critical. Very few shocks are in the endogenous critical class.

Measurement

Articles that experience shocks can differ in various collaborative measures that potentially affect newcomer participation (e.g., level of activity, feedback). In this section, we formally define these measures as well as measures of participation. We consider each article as a unit and the editors who edit an article collectively as a collaborative crowd [2]. The measures are based on the weekly level and the week in which the shock takes places is indexed as week 0.

Level of activity. For an article in week t, let E_t be the set of editors and W_t be the set of edits. We measure the level of activity by $|E_t|$ and $|W_t|$. We use superscript NC and IC to distinguish between article newcomers [3] (i.e., E_t^{NC} refers to those who edit the article for the first time in week t) and incumbents (i.e., E_t^{IC} refers to those who edit the article prior to week t as well as at t).

Centralization. Similar to related work (Zhang et al. 2017a; 2017b), we measure centralization of workload in a group using the *normalized Gini coefficient* of the distribution of number of revisions contributed by each editor. The Gini coefficient measures the level of inequality in a distribution (Dorfman 1979). Normalized gini coefficient accounts for heterogeneity in group sizes by normalizing the Gini coefficient by the maximum possible Gini coefficient given the number of editors and revisions. Therefore, it measures the level of centralization that is driven by how workload is internally distributed among editors. For a group with an editor set E_t and edits set W_t, the centralization for an article during week t is defined as:

$$C_t = \frac{\sum_{i \in E_t} \sum_{j \in E_t} |e_i - e_j|}{2(|E_t| - 1)(|W_t| - |E_t|)}$$

where e_i and e_j represent the number of edits by editor i and j during week t to an article.

Negative feedback to newcomers. Wikipedia has a reverting feature that allows editors to undo revisions by other editors. In order to measure the negative feedback received by a group of editors E' during week t, we use the fraction of their edits during that week being reverted eventually. We denote this measure as $R^{i,j}_{E'}$. We are particularly interested in the negative feedback received by newcomers ($R^{i,j}_{E_t^{NC}}$).

While we interpret reverts as negative feedback, other studies have associated reverts with conflict and controversy

[2] We exclude bots using: https://github.com/halfak/are-the-bots-really-fighting/blob/master/datasets/crosswiki_category_bot_20170328.tsv

[3] For the rest of the paper we refer to "article newcomers" simply as "newcomers" for brevity.

(Kittur et al. 2007; Kittur and Kraut 2010; Zhang et al. 2017b; 2017a). One key difference is that we focus on reverts to article newcomers, who due to their lack of expertise, are more likely to learn about the norms and rules of the article by being reverted than other more seasoned editors.

Another concern is that some reverts can be due to the changing nature of shock events. Certain edits can become stale as events progress.

We collect a random sample of 100 reverts of edits by newcomers with a comment explaining the reason for reverting. We find that only 4% are consistent with reverts due to uncertainty in information as the story was unfolding. Thus, while we find that some reverts are responding to fluid information, the majority of them are due to other reasons such as lack of citations and lack of compliance to Wikipedia rules.

Discussion.

The talk page of each article provides a forum for coordinating efforts and opinions. The amount of discussions can potentially affect the future participation of newcomers by providing them with guidance. We measure discussion by the number of comments on the talk page $|D_t|$ where D_t is the set of comments on the talk page in week t. In addition to total number of comments, we also measure (i) the log ratio of comments to revisions $\log(|D_t|+1)/\log(|W_t|+1)$ to characterize how much discussion takes place per contribution, and (ii) discussions per newcomers (D_t^{NC}) to identify how much newcomers engage in discussion.

Future Group Participation. There are several possible ways to measure participation. We choose a measure that reflects the dynamics on Wikipedia during times of shocks. In essence, any measure of participation should capture *whether* and *how much* a particular user or set of users continue to contribute to an article. However, given the fluctuations in the amount of work that an article needs during a shock, we must also consider the availability of work when measuring participation. Since shocks often lead to article update needs, and thus attract editors, we cannot simply measure the fraction of users who returned to the article or the number of edits they make. This measure would be artificially inflated for those who arrived immediately before the shock and artificially deflated for those who arrived at the shock. Instead, we measure the number of users who return to the article, while controlling for the total number of active users, which serves as a proxy for the amount of work available in the article. Based on this principle and in order to ensure that our results are not dependent on a particular way of measuring participation, we develop three different measures of future participation.

1. Unweighted Future Participation: For any article, we measure the *unweighted future participation* of a subset of users during week t, E'_t, as the fraction of active editors in weeks $t+1$ — $t+4$ who are also part of the set E'_t. That is, unweighted future participation of users E'_t is $\frac{|E_{[t+1,t+4]} \cap E'_t|}{|E_{[t+1,t+4]}|}$, where $E_{[t+1,t+4]}$ is the set of users who were active during weeks $t+1$ — $t+4$. This measures what fraction of the workforce users in E'_t accounted for during weeks $t+1$ — $t+4$.

2. Weighted Future Participation: An important consideration is that there is high variance in the level of contribution among users. Thus, we can also weight participation by the actual number of revisions that users contributed, not just by whether they were active. We measure the *weighted future participation* of a subset of active users during week t, E'_t, as the fraction of revisions to the article during weeks $t+1$ — $t+4$ that were contributed by editors in E'_t.

3. Sustained Participation: We aim to capture how E'_t *changed* their level of contribution from week t to weeks $t+1$ through $t+4$. Thus, we define the *sustained participation* of a subset of active users during week t, E'_t, as the difference in their fractional contributions during weeks $t+1$ — $t+4$ and their fractional contributions during week t. The sustained participation of E'_t is $\frac{|W^{E'_t}_{[t+1,t+4]}|}{|W_{[t+1,t+4]}|} - \frac{|W^{E'_t}_t|}{|W_t|}$,

where $W^{E'_t}_T$ and W_T are the set of revisions by editors E'_t and by all editors during time interval T, respectively.

Our definitions of participation operate at the group level rather than the individual level. We are interested in how shocks impact future participation of different groups of editors, particularly newcomers and incumbents. Throughout the paper, we report results using the weighted and unweighted future participation. Sustained participation results are qualitatively similar and are omitted for brevity.

Results

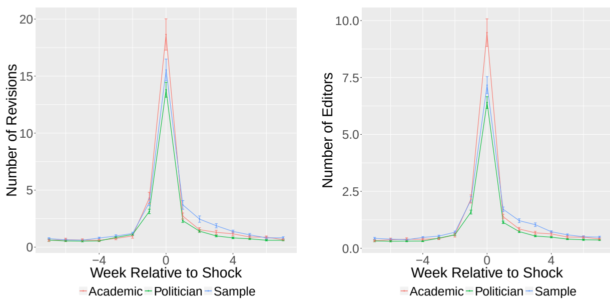
Change in Collaboration Dynamics with Shocks

We track the changes of each collaboration feature eight weeks before and after the shock. We produce time series that show the mean value of each feature versus the number of weeks relative to the shock ($t = 0$). We observe that some features begin changing before the week of the shock. This is expected since the shock was defined as the time when attention peaked, not when the shock occurred. We now describe how the shock affects each collaboration measure.

Level of Activity. Figure 2 illustrates the amount of activity, as measured by the number of editors and edits to the article. The solid lines represent the average value and the error bars denote the standard error. The three curves show the results for the different types of articles: academics, politicians, and the sample of Biography articles. The same convention is followed in most of the rest of the figures. As expected, the amount of activity increases sharply around the time of the shock. This further validates our shock identification procedure. Indeed, Wikipedia and search engines (such as Google) play an important role in each other's traffic and quality (McMahon, Johnson, and Hecht 2017).

We measure the fraction of first time editors to the article on a week t as ($|E_t^{NC}|/|E_t|$). The fraction of article newcomers increases during the shock (Figure 4(a)) suggesting that attention shocks provide an opportunity to recruit new editors to the article.

We consider the newcomers to articles during the week of the shock ($t = 0$) and compare their experience editing other articles in Wikipedia prior to $t = 0$ with that of their incumbent counterparts. Specifically, we compare their number of edits across Wikipedia (excluding the edits by the incumbents to the focal article that had the attention shock). Figure 3 shows the distribution of the number of edits across Wikipedia by newcomers and incumbents. We find that newcomers tend to have significantly less experience in Wikipedia than incumbents. The mean number of edits by incumbents is 21,673 while that of newcomers is 10,981 (a t-test confirms that this difference is significant $p < 0.001$). A non-parametric bootstrap test (Ghosh et al. 1984) also indicates that the median number of edits by newcomers (2,617) is significantly lower ($p < 0.014$) than that of incumbents (4,491). This shows that while newcomers are not new to Wikipedia, they do come to articles with far less Wikipedia experience than incumbents, suggesting that they can benefit for rich interactions with incumbents such as discussion in talk pages in order to become more familiar with the norms and policies of the article and Wikipedia in general.



(a) Number of Edits (b) Number of Editors

Figure 2: Volume of Article Activity

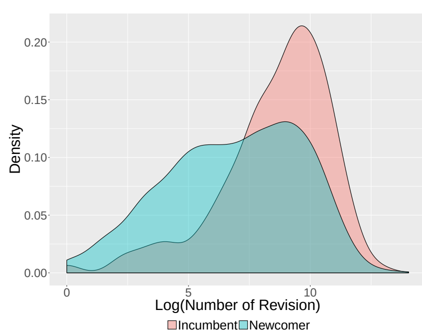
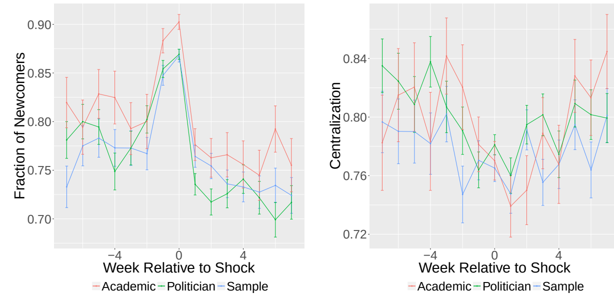


Figure 3: Distribution of edits across all Wikipedia articles by newcomers and incumbents prior to the day of the shock. Edits by incumbents to the shocked article are excluded from their count.

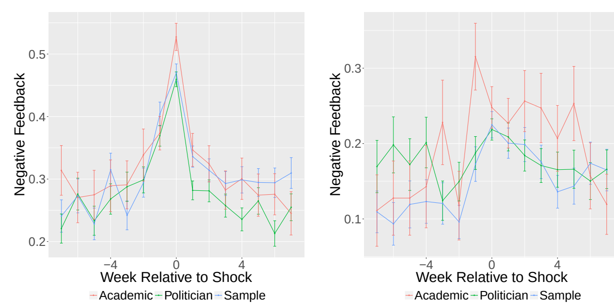
Centralization. Centralization tends to decrease starting three weeks before the shock and returns to its regular level by $t = 5$ (Figure 4(b)). This is consistent with the intuition that the events triggering the shock increase (i) the amount of content that needs to be added to the article, and (ii) the number of people willing to contribute. Thus, editors who normally would not take on a large workload now contribute more and make the workload more evenly distributed.



(a) (b)

Figure 4: (a) Fraction of Newcomers, (b) Centralization

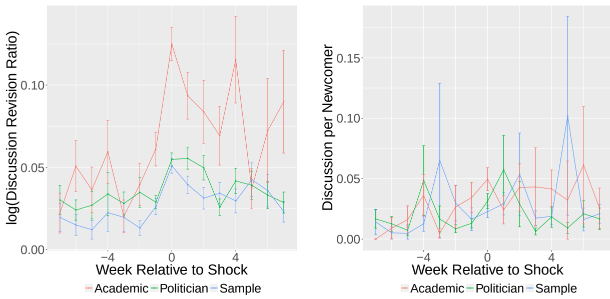
Negative Feedback to Newcomers. During the shock, articles are more likely to direct negative feedback in the form of reverts towards newcomers. The rate of negative feedback toward incumbents also increases slightly, though there is larger variance and the magnitude of the effect is much smaller compared to the newcomers (Figure 5). Newcomers, having less experience with the norms of the article, are generally reverted more than other users. Furthermore, given that the shock is likely to be a chaotic period for the article, it is reasonable that we observe more overall reverting.



(a) Newcomers (b) Incumbents

Figure 5: Fraction of articles that had negative feedback directed towards newcomers and incumbents

Discussion. We may expect editors to engage less on discussion and focus more on updating the main article given the timeliness of shocks. On the other hand, the shock may lead to more complexity, requiring coordination by the editors and hence more discussion. We observe a slight increase in the amount of discussion around the time of the shock, suggesting that the latter hypothesis is better aligned with the dynamics on Wikipedia (figure 6(a)). We also track the extent to which newcomers engage with the discussion page. However, we observe no changes in the number of discussion edits per newcomer when the shock occurs.

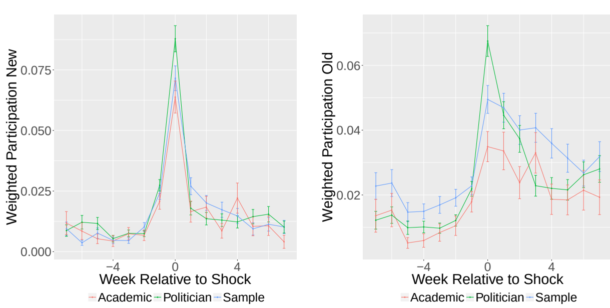


(a) Log ratio of edits on talk page to edits on article

(b) Fraction of edits on talk page per newcomer

Figure 6: Talk Page Activity

Future Participation. We now explore how the future participation editors changes as a function of time, relative to when the shock occurs. Recall that our future participation measures at a time t are based on activity during weeks $t+1$ through $t+4$. While weighted future participation is always low ($< 7\%$), it tends to peak for both newcomers and incumbents at the time of the shock. Future Participation of newcomers relaxes back to its non-shock baseline after the peak of the shock (Figure 7). Yet, incumbent future participation relaxes much slower and remains significantly higher than non-shock baseline even after eight weeks. This suggests that shocks make articles better at maintaining editors engaged. However, the effect is not as strong or sustained for the newcomers. Figure 8 shows that the results are similar for unweighted future participation. Next, we will investigate the factors that are associated with a higher future participation of newcomers.



(a) Newcomers

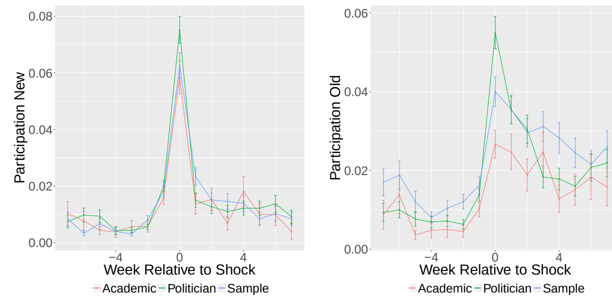
(b) Incumbents

Figure 7: Weighted Future Participation

Mediation Analysis

Our aim is to go beyond analyzing how participation changes and further explain what mechanisms, i.e. changes in the aforementioned collaboration dynamics, result in such change. To that end, we perform mediation analysis which aims to uncover the underlying mechanism by which one variable influences another variable through one or more mediator variables (MacKinnon, Fairchild, and Fritz 2007).

Our analysis focuses on a timeframe that is within five



(a) Newcomers

(b) Incumbents

Figure 8: Unweighted Future Participation

weeks relative to the shock. We perform a before-after comparison by encoding 2 weeks prior to the shock as the non-shock period and 3 weeks after as the shock period. We first report the results using the two-week timeframe. We choose this small time frame so that the results are minimally affected by the general trends over Wikipedia. We also perform our analysis on various choices for non-shock/shock periods and report robustness checks.

We use a structural equation model (Baron and Kenny 1986) to estimate our mediation model (see Figure 1):

$$\begin{aligned}
&\text{future participation} = b_0 \times \text{shocks} \\
&+ b_1 \times \text{centralization} + b_2 \times \text{negative feedback to NC} \\
&+ b_3 \times \text{total discussion} + b_4 \times \text{discussion per NC} \\
&+ b_5 \times \text{article size} + b_6 \times \text{fraction of NC} \\
&+ b_7 \times \text{fraction of negative feedback} \qquad (1)
\end{aligned}$$

and mediation equations:

$$\begin{aligned}
\text{centralization} &= a_1 \times \text{shocks} \\
\text{negative feedback to NC} &= a_2 \times \text{shocks} \\
\text{total discussion} &= a_3 \times \text{shocks} \\
\text{discussion per NC} &= a_4 \times \text{shocks} \qquad (2)
\end{aligned}$$

Under this specification, b_0 captures the direct impact of the shock on newcomer participation. The model also incorporates four indirect channels through which the shock exerts impact: 1) centralization, 2) negative feedback to newcomers, 3) total discussion on the talk page, 4) average number of comments by newcomers. For each channel, b_i captures how the measure affects the participation of newcomers and a_i captures how the measure is affected by the shock. The indirect effect from the shock through the mediators is measured by $a_i \times b_i$. We include the article size, fraction of negative feedback, and fraction of newcomers as controls. Effects estimated using the delta method (Sobel 1982).

Table 5 shows how the four mediators are affected by the shock (captured by a_1 through a_4 in Equation 2). Centralization is not statistically significantly related to the participation of newcomers. Both negative comments to newcomers and the amount of discussions on the talk page have significantly positive relationship with the participation of newcomers. Negative feedback to newcomers and the amount

	Weighted		Unweighted	
	Coeff.	SE	Coeff.	SE
shocks	0.011^{**}	0.004	0.013^{***}	0.003
centralization	0.005	0.004	0.003	0.004
neg. feedback to NC	0.068^{***}	0.005	0.049^{***}	0.004
total discussion	0.001^{***}	0.000	0.001^{***}	0.000
discussion per NC	0.002	0.005	0.003	0.004
article size	-0.021^{***}	0.001	-0.018^{**}	0.001
frac. NC	0.034^{***}	0.009	0.029^{**}	0.008
frac. neg. feedback	-0.118^{***}	0.011	-0.085^{**}	0.010
constant	0.086^{***}	0.011	0.072^{**}	0.009

Table 4: Regression results (coefficient and standard error) from Equation 1 for weighted and unweighted future participation. (***: p-val$\leq$ 0.001, **: p-val $\leq$ 0.01, *: p-val $\leq$ 0.05).

	Coeff.	SE
centralization	-0.019^{*}	0.009
neg. feedback to NC	0.136^{***}	0.012
total discussion	0.991^{***}	0.228
discussion per NC	0.005	0.009

Table 5: Regression results from the mediation equations 2

of discussions on the talk page increase significantly and centralization decreases significantly. These results are all in line with findings presented in Section .

Tables 4 through 6 show the results from the mediation analysis. Table 4 reports the estimation results for the main equation. For each measure, we report the parameter estimate (b_0 through b_7 in Equation 1) and the standard error. The shock has a direct effect of 1.128 percentage points on the participation of newcomers. Compared to the baseline participation of 8.639 percentage points, the effect size of the shock directly (1 percentage point) amounts to a 13% increase in participation. Our results also hold for unweighted future participation (see Table 6). Compared to the baseline future participation of 7.225 percentage points, the effect size of the shock directly is 1.323 percentage points (18%).

Table 6 summarizes the direct effect and indirect effects from the four channels. The results show that the total effect that the shock exerts on the participation of newcomers is 2.173 percentage points. That is, after the shock, the participation of newcomers nearly doubles. Out of the overall impact, approximately 52%(1.128 percentage point) is direct and 48% (1.048 percentage point) is indirect through negative feedback to newcomers.

Robustness Test - Time frame We conduct mediation analysis over various time frame choices. Specifically, we fix the start of the non-shock period five weeks prior to the shock, and vary the end between three weeks prior to the shock and one week prior to the shock. The start of the shock period follows immediately after the non-shock period and the end varies between one week and five weeks after the shock. For the direct effect and each indirect effect, we report the mean estimates calculated from each model as well as the minimum and maximum in Table 7. We find that the direction and magnitude of the effect size across time frame choices are consistent.

	Weighted		Unweighted	
	Coeff.	SE	Coeff.	SE
Indirect Effect	1.045^{***}	0.1.08	0.746^{***}	0.018
centralization	-0.010	0.000	-0.005	0.008
neg. feedback to NC	0.919^{***}	0.103	0.666^{***}	0.080
total discussion	0.135^{***}	0.036	0.084^{***}	0.025
discussion per NC	0.008	0.003	0.002	0.003
Direct Effect	1.128^{**}	0.400	1.323^{***}	0.343
Total Effect	2.173^{***}	0.403	2.068^{***}	0.343
Baseline Part.	8.639		7.225	

Table 6: Effects from mediation analysis

	Average	Min	Max
Indirect Effect			
centralization	-0.010	-0.010	-0.010
negative feedback to NC	0.800	0.500	1.100
total discussion	0.134	0.101	0.138
discussion per NC	0.000	0.000	0.000
Direct Effect	1.124	1.120	1.128
Total Effect	2.170	2.169	2.173

Table 7: Mediation analysis results with various time frames

Robustness Test - Article Size One might also be concerned that the effects of the shock might be heterogeneous in the size of the articles. To test this, we perform the mediation analysis separately for the small and large articles. We define small (or large) articles as the ones whose size at the time of the shock is below (above) the median of the empirical distribution. Procedurally, articles with no more than 35 revisions are classified as small articles. Table 8 summarizes the mediation analysis based on article size. We find that the signs of the effect are preserved between the small and large articles, while the magnitude varies. The baseline participation is 9.39 percentage points for the small articles and 5.95 percentage points for the large articles. Given that, the total effect from the shock is 3.142 percentage points (33% of the baseline) for small articles and 2.307 percentage points (39% of the baseline) for large articles. Among the small articles, the direct effect from the shock on participation is 2.410 percentage points, which is more than two times larger than the indirect effect of 1.011 percentage points. In contrast, among the large articles, the direct effect and the indirect effect are similar in the magnitude. While there are some differences in magnitude and fraction of the total effect explained by indirect effects, the signs of the effect are consistent. Overall, the results of the mediation model are qualitatively similar between the two groups of articles.

	Small		Large	
	Coeff.	SE	Coeff.	SE
Indirect Effect (p.p)	1.011^{***}	0.233	1.052^{**}	0.366
Direct Effect (p.p)	2.410^{*}	1.129	1.254^{***}	0.123
Total Effect (p.p)	3.142^{**}	1.132	2.307^{***}	0.370
Baseline Part. (p.p)	9.388^{**}	2.792	5.949^{***}	1.171

Table 8: Mediation results for Small and Large Articles

Contribution Spillover

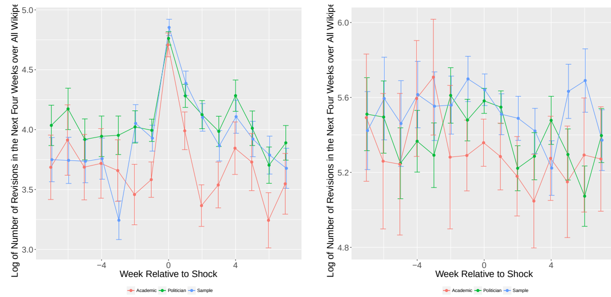
Our results establish that shocks have an impact on articles that experience a shock. Next, we investigate whether the shock has spillover effects on the contributions to *other* articles across Wikipedia. To address this question, for each editor who is active on an article during a given week, we define the editor's *spillover participation* as the editor's number of revisions over all articles (excluding the focal one) in the next four weeks and apply log-transformation to mitigate the effect from outliers. For each article, the spillover effect is the average spillover participation of its editors. Our results establish that shocks have an impact on articles that experience a shock. An open question is whether the shock has spillover effects on the contributions to *other* articles across Wikipedia. To address this question, for each editor who is active on an article during the week of a shock, we define the editor's *spillover participation* as the editor's number of revisions over all articles (excluding the focal one) in the next four weeks. To mitigate the effect from outliers, we apply the log-transformation to the spillover participation of each user. For each article with a shock, the spillover effect is the average spillover participation of its editors.

Figures 9(a) and 9(b) show the spillover participation for the newcomers and the incumbents of the focal articles. The shock has a remarkable impact on the spillover of newcomers – compared to the non-shock weeks, the shock yields higher spillover over all Wikipedia for newcomers. A series of t-tests indicate that spillover is significantly larger during the week of the shock ($p- < 0.01$ for all comparisons). In contrast, we see from figure 9(b) that the shock does not lead to more contributions in other articles by incumbents. This suggests that when shocks attract newcomers to an article, they also produce an increase their contributions to other articles across Wikipedia. This further highlights the importance of understanding the process by which attention shocks bring new editors to an article and keep them engaged.

Summary of Findings

Our results can be organized into three major findings. First, on the relationship between shocks and collaborative measures, our results show support for the relationship between shocks and centralization (H1, supported), reverts of newcomers (H2, supported), edits to talkpages (H3, supported) but not newcomer's edits to talkpages (H4, not supported).

Second, on the relationship between collaborative measures and future participation, our results find support for the relationship between reverts of newcomers (H6, supported)



(a) Newcomers (b) Incumbents

Figure 9: Log-transformation of number of revisions on all Wikipedia but the focal article

and edits to talkpages (H7, supported).However, our results did not find a relationship between future participation and centralization (H5, not supported) nor with newcomer's edits to talkpages (H8, not supported).

Finally, we find support for the mediation effects of shocks through the collaborative mechanisms (H9). The relationship between shocks and future participation is partially mediated by reverts of newcomers (supported) and edits to talkpages (supported). Furthermore, our mediation results are consistent across articles of different size.

Discussion

Our goal in this research is to examine the impact of attention shocks on crowd participation and to identify factors that mediate this effect. In doing so, this study provides a number of new insights into how shocks can facilitate participation in public good contributions.

First, our research expands the current understanding of public good contributions by highlighting how attention shocks lead to increased crowd participation. Many articles on Wikipedia suffer from a lack of interest from editors. Attention shocks offer an opportunity to attract editors to an article and promote their continued and increased participation. Yet, most previous research has not examined the impact of shocks. Our study fills this gap and suggests that shocks can be leveraged to help encourage editors contribute to Wikipedia articles they have not contributed to in the past.

Second, our findings contribute to the cumulative research on the impacts of reverts on Wikipedia participation. While some studies suggest that negative feedback reduces participation (Hausknecht and Trevor 2011; Halfaker et al. 2013), others suggest they can lead to an increase (Zhu et al. 2013a). We find that reverts of newcomers is associated with increased future participation. This finding contributes to the literature of conflicting claims and highlights that the relationship between negative feedback and participation is context dependent. Note that our study does not examine newcomers to a platform but instead newcomers to the article. New editors to articles, relative to new editors of Wikipedia, may be more comfortable having their edits reverted and may actually learn about the article's community through reverts. Newcomers to an article may view reverts as helpful feedback rather than conflict, which is consistent with prior work on the positive effects of negative, directive feedback on work output in Wikipedia (Zhu et al. 2013a). Furthermore, the positive relationship observed might be due to content—editors might be providing better explanations in the revert edit comments for attention shock articles compared to other cases of reverting. Future work that delves deeper into such analysis can provide further insights.

Third, this study highlights the importance of discussion on talkpages on the participation of crowd members following shocks. An increase in edits to talk pages as a result of the shock leads to increased future participation of newcomers to the article. Edits to talk pages appear to be an indicator of community discussion and coordination around the events leading to the shocks. These discussions may help to socialize newcomers and create a sense of community. Ultimately, the need to come together and react to the shock may be the most valuable aid to increased participation.

Finally, our results show that studying mediating mechanisms provides a more informed understanding of how attention shocks lead to increased future crowd participation. We find evidence that attention shocks affect future participation in different ways, via reverts of newcomers and discussions on talkpages. These results help to explain not only how attention shocks affect participation but highlights the *mechanisms* through which such effect takes place.

Design implications. Findings from this study can inform the design of the Wikipedia platform features to encourage further participation from newcomers to articles. Our study shows that directly engaging newcomers to an article in the form of reverts leads to greater future participation on the focal article. To leverage this finding, we suggest adding features that can help facilitate this type of interaction. For example, the platform can automatically track and flag at least one contribution from a crowd newcomer for review. These flags can act as nudges to remind existing crowd members to provide feedback to newcomers.

Another important design implication relates to the benefit of discussions on talkpages. While other studies have suggested the need to support discussion on Wikipedia, our findings suggest that such features may be particularly valuable during times of attention shocks. The need to include more users in the discussion in a short time period may require a partial redesign of the current talkpages.

We hope that our findings generate useful insights and further research to begin viewing exogenous events that generate traffic to Wikipedia articles as "happy accidents" that, given the appropriate platform design, can bring new contributors and more diverse perspectives rather than chaotic periods that end with locked articles and potentially interested editors without the ability to contribute their expertise.

We also note that the findings presented in this study have potential implications for other collaborative crowdsourcing platforms (e.g. GitHub). Indeed, collaboration dynamics studied here are general and can be measured for other platforms as well. Similarly, comparable shocks can be identified for other platforms. Future research that investigates the relationship between such shocks and collaboration dynamics will help solidify the generalizability of our findings.

Limitations and future research. Google Trends is a rich source of data and we believe our methodology can be applied in broader settings to identify attention spikes to other topics and study shocks on online communities beyond Wikipedia. However, our approach has limitations.

According to our validation, the majority of such spikes correspond to real events. However, we were unable to validate a small but non-negligible fraction of the shocks. This suggests that there is noise in our data. Based on the clarity of some of the results, we believe that our approach is accurate enough to provide signal and answer our research questions. However, researchers using our approach should be aware of false positives. Additionally, a promising research direction is to explore more accurate ways of extracting shocks from search engine data such as Google Trends.

Finally, we focused on attention shocks directed towards people. While the effects did not vary with the type of person (academic, politician, or other), it is possible that the effects could be different for articles on different subjects such as organizations, locations, and history.

Acknowledgements

This research was partly supported by the National Science Foundation under Grant No. IIS-1617820. We thank Brian Keegan, Aaron Halfaker, Paul Resnick, and Nazanin Andalibi for helpful discussions and feedback. We also thank Aaron Halfaker for providing a list of Wikipedia bots.

References

Antin, J., and Cheshire, C. 2010. Readers are not free-riders: Reading as a form of participation on wikipedia. In *CSCW*, CSCW '10, 127–130. New York, NY, USA: ACM.

Arazy, O., and Nov, O. 2010. Determinants of wikipedia quality: the roles of global and local contribution inequality. In *CSCW*, 233–236. ACM.

Baron, R. M., and Kenny, D. A. 1986. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology* 51(6):1173.

Buechler, S. M. 2013. Strain and breakdown theories. *The Wiley-Blackwell Encyclopedia of Social and Political Movements*.

Butler, B. S. 2001. Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information systems research* 12(4):346–362.

Chen, J.; Ren, Y.; and Riedl, J. 2010. The effects of diversity on group productivity and member withdrawal in online volunteer groups. In *CHI*, 821–830. ACM.

Choi, B.; Alexander, K.; Kraut, R. E.; and Levine, J. M. 2010a. Socialization tactics in wikipedia and their effects. In *CSCW*, CSCW '10, 107–116. New York, NY, USA: ACM.

Choi, B.; Alexander, K.; Kraut, R. E.; and Levine, J. M. 2010b. Socialization tactics in wikipedia and their effects. In *CSCW*, 107–116. ACM.

Ciampaglia, G. L., and Taraborelli, D. 2015. Moodbar: Increasing new user retention in wikipedia through lightweight socialization. In *CSCW*, 734–742. ACM.

Cleveland, R. B.; Cleveland, W. S.; and Terpenning, I. 1990. Stl: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* 6(1):3.

Crane, R., and Sornette, D. 2008. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences* 105(41):15649–15653.

Dorfman, R. 1979. A formula for the gini coefficient. *The Review of Economics and Statistics* 146–149.

Faulkner, R.; Walling, S.; and Pinchuk, M. 2012. Etiquette in wikipedia: weening new editors into productive ones. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*, 5. ACM.

Ghosh, M.; Parr, W. C.; Singh, K.; Babu, G. J.; et al. 1984. A note on bootstrapping the sample median. *The Annals of Statistics* 12(3):1130–1135.

Halfaker, A.; Geiger, R. S.; Morgan, J. T.; and Riedl, J. 2013. The rise and decline of an open collaboration system: How wikipedias reaction to popularity is causing its decline. *American Behavioral Scientist* 57(5):664–688.

Halfaker, A.; Kittur, A.; and Riedl, J. 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*, 163–172. ACM.

Hausknecht, J. P., and Trevor, C. O. 2011. Collective turnover at the group, unit, and organizational levels: Evidence, issues, and implications. *Journal of Management* 37(1):352–388.

Jones, J. 2008. Patterns of revision in online writing: A study of wikipedia's featured articles. *Written Communication* 25(2):262–289.

Karumur, R. P.; Yu, B.; Zhu, H.; and Konstan, J. A. 2018. Content is king, leadership lags: Effects of prior experience on newcomer retention and productivity in online production groups. In *CHI*, 506. ACM.

Keegan, B.; Gergle, D.; and Contractor, N. 2012. Do editors or articles drive collaboration?: multilevel statistical network analysis of wikipedia coauthorship. In *CSCW*, 427–436. ACM.

Keegan, B.; Gergle, D.; and Contractor, N. 2013. Hot off the wiki: Structures and dynamics of wikipedias coverage of breaking news events. *American Behavioral Scientist* 57(5):595–622.

Kittur, A., and Kraut, R. E. 2008. Harnessing the wisdom of crowds in wikipedia: Quality through coordination. In *CSCW*, 37–46. ACM.

Kittur, A., and Kraut, R. E. 2010. Beyond wikipedia: coordination and conflict in online production groups. In *CSCW*, 215–224. ACM.

Kittur, A.; Suh, B.; Pendleton, B. A.; and Chi, E. H. 2007. He says, she says: conflict and coordination in wikipedia. In *CHI*, 453–462. ACM.

Krieger, M.; Stark, E. M.; and Klemmer, S. R. 2009. Coordinating tasks on the commons: designing for personal goals, expertise and serendipity. In *CHI*, 1485–1494. ACM.

Leung, W.; Zhu, H.; and Konstan, J. A. 2017. The effect of emotional cues from the nfl on wikipedia contributions. *Proc. ACM Hum.-Comput. Interact.* 1(CSCW):66:1–66:21.

Li, A., and Farzan, R. 2018. Keeping up on current events! a case study of newcomers to wikipedia. In *International Conference on Social Informatics*, 348–369. Springer.

MacKinnon, D. P.; Fairchild, A. J.; and Fritz, M. S. 2007. Mediation analysis. *Annual Review of Psychology* 58:593.

McMahon, C.; Johnson, I.; and Hecht, B. 2017. The substantial interdependence of wikipedia and google: a case study on the relationship between peer production communities and information technologies. In *ICWSM*. AAAI Press.

Mesgari, M.; Okoli, C.; Mehdi, M.; Nielsen, F. Å.; and Lanamäki, A. 2015. "the sum of all human knowledge": A systematic review of scholarly research on the content of wikipedia. *Journal of the Association for Information Science and Technology* 66(2):219–245.

Morgan, J. T., and Halfaker, A. 2018. Evaluating the impact of the wikipedia teahouse on newcomer socialization and retention. In *Proceedings of the 14th International Symposium on Open Collaboration*, 20. ACM.

Platt, E., and Romero, D. 2018. Network Structure, Efficiency, and Performance in WikiProjects. In *Proceedings of International AAAI Conference on Web and Social Media (ICWSM)*.

Ransbotham, S., and Kane, G. C. 2011. Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in wikipedia. *Mis Quarterly* 613–627.

Robert, L., and Romero, D. M. 2015. Crowd size, diversity and performance. In *CHI*, 1379–1382. ACM.

Robert, L. P., and Romero, D. M. 2017. The influence of diversity and experience on the effects of crowd size. *Journal of the Association for Information Science and Technology* 68(2):321–332.

Romero, D. M.; Huttenlocher, D.; and Kleinberg, J. 2015. Coordination and efficiency in decentralized collaboration. In *ICWSM*.

Schneider, J.; Gelley, B. S.; and Halfaker, A. 2014. Accept, decline, postpone: How newcomer productivity is reduced in english wikipedia by pre-publication review. In *Proceedings of the international symposium on open collaboration*, 26. ACM.

Sobel, M. E. 1982. Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology* 13:290–312.

Staw, B. M.; Sandelands, L. E.; and Dutton, J. E. 1981. Threat rigidity effects in organizational behavior: A multilevel analysis. *Administrative Science Quarterly* 501–524.

Suh, B.; Convertino, G.; Chi, E. H.; and Pirolli, P. 2009. The singularity is not near: slowing growth of wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*, 8. ACM.

Useem, B. 1998. Breakdown theories of collective action. *Annual Review of Sociology* 24(1):215–238.

Voß, J. 2005. Measuring wikipedia. In *Proceedings of 10th International Conference of the International Society forScientometrics and Informetrics, (Stockholm, Sweden)*.

Yu, B.; Wang, X.; Lin, A. Y.; Ren, Y.; Terveen, L. G.; and Zhu, H. 2017. Out with the old, in with the new?: Unpacking member turnover in online production groups. *PACMHCI* 1(CSCW):117–1.

Zhang, A. F.; Danielle, L.; Budak, C.; Robert Jr., L. P.; and Romero, D. M. 2017a. Shocking the crowd: The effect of censorship shocks on chinese wikipedia. In *ICWSM*.

Zhang, A. F.; Livneh, D.; Budak, C.; Robert, Jr., L. P.; and Romero, D. M. 2017b. Crowd development: The interplay between crowd evaluation and collaborative dynamics in wikipedia. *Proc. ACM Hum.-Comput. Interact.* 1(CSCW):119:1–119:21.

Zhu, H.; Zhang, A.; He, J.; Kraut, R. E.; and Kittur, A. 2013a. Effects of peer feedback on contribution: A field experiment in wikipedia. In *CHI*, CHI '13, 2253–2262. New York, NY, USA: ACM.

Zhu, H.; Zhang, A.; He, J.; Kraut, R. E.; and Kittur, A. 2013b. Effects of peer feedback on contribution: a field experiment in wikipedia. In *CHI*, 2253–2262. ACM.