WIKI HUES: HINDI, URDU, AND ENGLISH TAKES ON EVOLVING REGIONAL CONFLICT

HARSH SANGHAVI, MOLLY HICKMAN, AND VIRAL PASAD

Wikipedia at its core is the product of thousands of editors working collaboratively to provide free and up-to-date knowledge to anyone who uses their services. For Wikipedia to be truly accessible to anyone, it has to be accessible in different languages and has to ensure consistency in the information being presented in all its versions. The recent abrogation of Article 370 in the Constitution of India, as well as other recent events in and concerning Jammu and Kashmir, have spurred attention to related articles on Wikipedia–specifically, the English, Hindi and Urdu versions, both by information seekers and editors. This article compares the collective work of editors of the Hindi, Urdu and English versions of Wikipedia articles related to the current conflict in Jammu and Kashmir. Additionally, this article asks to what degree the articles in each language achieve Wikipedia's mission of making neutrally-presented, reliable information available to people around the globe. To answer these questions, correlations between the volume, nature, and timing of contributions by editors following recent events were investigated for Hindi, Urdu and English Wikipedia articles. Time series analysis of edit data showed that in all three languages, edits closely followed significant events, with editors in all three languages responding within two days of increased traffic to a given page. Semi-structured interviews of editors from all three versions of Wikipedia showed a universal focus on the neutral point of view, while also shedding light on the collaborative nature of Wikipedia editing and the response protocols editors follow after additions of misinformation and spam by vandals. Finally, editor perceptions on the nature of collaborative work and conflicts was investigated, revealing the close knit nature of Hindi and Urdu editors in their editing process.

1 INTRODUCTION

Wikipedia defines itself as a multilingual online encyclopedia, created and maintained as an open collaboration project using a wiki-based editing system. Its mission is to encourage consensus around a single neutral point of view. [16] However one of the difficulties in achieving that mission has been due to Wikipedia Foundation setting up editions in over 250 different languages. This means that every subject matter (referred to as 'topic') could have pages ('articles') in those 250 editions. There are many language editions of Wikipedia, the most frequently edited and visited of which is, unsurprisingly, the English edition[14] Lesser-edited languages see fewer and generally shorter articles than English. The lack of sharing of information between different language editions of Wikipedia prevents access to a larger variety of content for mono-lingual users. [16]

The information gap problem is exacerbated when the article is related to a polarizing topic. Information-seekers who are in proximity to a conflict, and only refer to articles on said conflict in their local language, may get an incomplete picture of the events. We are interested in how Wikipedia editors in two lesser-edited languages, Hindi and Urdu, create a collective memory around the long-standing and evolving conflict in Jammu and Kashmir, compared to one another and to the editors of the corresponding articles in English.

A brief snapshot: in the English version of *Kashmir Conflict* page, there is a section entitled 'Water dispute'. In April 2017 an Urdu contributor appears to have seen the English version and added such a section to the Urdu article, as well as several at least one other section that was then, but is no longer, on the English counterpart. This edit more than doubled the size of the Urdu article; however, within a few months, another user had undone the edit, without much in the way of an explanation. The Hindi counterpart also referred to the water dispute at one time, but it was also redacted. We ask ourselves, why? What events may have precipitated the addition and subsequent removal of this section? What motivated the adders and the removers?

1.1 Background:Jammu and Kashmir

Ever since gaining independence from the British, Pakistan and India have been in conflict over the Kashmir valley (Jammu Kashmir in India), a disputed region in the northern region of both countries where they border. While both countries only control a part of the former princely state, both claim Jammu and Kashmir (J&K) in its entirety [17]. Three quarters of the region's population is Muslim, which is one of Pakistan's (itself a Muslim majority Islamic Republic) claims to Kashmir among other factors. Three wars and various conflicts within the region have resulted in J&K divided between India, Pakistan and China. The line of control demarcates the regions administered by each nation. An insurgency began to proliferate in Indian-administered Kashmir in the late 1980s.

1.1.1 **Pulwama Attack of 2019**. According to Time, unrest in Kashmir grew in 2016 after India killed a popular militant leader, Burhan Wani [1]. On 14 February 2019, a convoy of vehicles carrying security personnel on the Jammu Srinagar National Highway was attacked by a vehicle-borne suicide bomber at Lethpora (near Awantipora) in the Pulwama district, Jammu and Kashmir, India. The attack resulted in the deaths of 40 Central Reserve Police Force (CRPF) personnel and the attacker. The responsibility for the attack was claimed by the Pakistan-based Islamist militant group Jaish-e-Mohammed. The attacker was Adil Ahmad Dar, a local from Pulwama district, and a member of Jaish-e-Mohammed.[9] India has blamed Pakistan for the attack. Pakistan condemned the attack and denied any connection to it.[11]

1.1.2 **Revocation of the constitutional Article 370**. Jammu and Kashmir (the parts administered by and also the parts claimed by India) had its own Constitution, thanks to the Article 370 of the Constitution of India. Hence, all laws passed by the Indian Parliament were not directly applicable to the state, unless the state government gave its concurrence. Along with Article 370, Article 35A gave citizens of J&K special status by barring outsiders from buying property and having Kashmiri women marrying non-Kashmiris lose their property rights of the State [24]. On the 5th of August 2019, the President of India issued an order overriding the prevailing order and nullifying all the provisions of autonomy granted to the state. Thereafter, the Home Minister of India introduced a reorganization bill in the Indian Parliament, to divide the state into two union territories.[6]

This situation has brought unrest into the region and kept peace between both India and Pakistan lying on a knife's edge. Consequently, as shown in previous research, such events give rise to multiple additions to and edits of Wikipedia articles pertaining to this issue [23]. Additionally, a majority of Pakistanis speak Urdu while a large percentage of Indians speak Hindi, and English is a major secondary language in both countries. The close proximity of both countries joined with their shared histories and political rivalries could possibly lead to conflicts in edits between these three languages. These factors contribute significantly to our motivation to study the Kashmir Conflict and its effects on Wikipedia edits.

2 RELATED WORK

2.1 Prior Work

Hecht et al[16] have explained how Wikipedia is a global source of world knowledge and how knowledge diversity can be exploited to create 'culturally aware' and hyper-lingual' applications. Their other work[15] highlights how self focus in community maintained knowledge repositories can be explained. Furthermore they express potential scope for analysis of self-focus in Wikipedia more thoroughly in controversial domains such as politics.

Hale[14] has examined roles of users editing multiple language editions of Wikipedia articles. His work shows that it might be wise to consider multilingual articles while considering design changes for Wikipedia. They discuss how multilingual editors bring over more context, sources and perspective from their main language articles to articles in other languages. Thus they reduce the level of self bias focus on Wikipedia. Hale[14] also show that multilingual editors often edit the same article in their primary and non-primary languages.

Along those same lines, Pfeil et al[20] have studied cultural differences between Wikipedia editors (specifically on the French, German, Japanese, and Dutch Wikipedias). They found several factors that differed between the cultures they studied, including reluctance to delete information, likelihood to collaborate with other editors, and propensity for spelling mistakes. We have not seen a similar study on any South Asian languages, but our work corroborates their finding of cultural differences in editing practices.

Keegan[19] have studied how Wikipedia editors collectively respond to breaking news events, and Twyman et al studied the same phenomenon as it related specifically to the Black Lives Matter movement [23]. New events, even if not directly related to the article at hand, will spur more page views, as well as edits that put the topic in the context of the new event: with every tragic new death of a black person at the hands of police, the *Black Lives Matter* page will see a spike in attention and edits, and often other articles in the web of articles related to the Black Lives Matter movement will also be viewed and edited [23].

Furthermore, Potts et al[12] have presented a study on the behavior and motivations of contributors on Health related articles on Wikipedia. They interviewed contributors and analysed the same using Thematic Analysis and a realist Grounded Theory approach. One of the key findings from their work is that contributors do not need to be experts on a given topic; lay people can also contribute to articles on Wikipedia, which is a theme that also echoes through our findings.

3 RESEARCH QUESTIONS

Our focus is on comparing Wikipedia editor behaviour, motivations and reactions to significant events related to the polarizing conflict in J&K, between the Hindi, Urdu and English versions of the site. With that in mind, we aim to answer the following research questions in our study.

(1) **RQ1: How do editors in Hindi, Urdu and English articles behave differently from each other immediately following significant events?**
   (a) After a significant event, the more editors an article has, the quicker the article will be updated in response to the event.
   (b) The Hindi and Urdu versions of a page will see vandalism and/or non-NPOV edits that become part of the stable version (i.e. persist, do not get redacted).
(2) **RQ2: How are editors and editing processes different across languages?**
   (a) For the same article, there will be significantly many common editors between languages.
   (b) For related topics, within a language, there will be significantly many common editors.
   (c) Collaboration and motivation will be different in the different languages. (open-ended)

4 DATA COLLECTION & METHODS

We drew data on revisions over the entire lifespan of the eight articles from the MediaWiki API[1], using Python[2]. The data, pulled on 10/29/2019, include the following fields:

- **user:** username or IP address
- **anon:** whether the user chose to be anonymous
- **timestamp:** date and time when the revision was made
- **size:** size in bytes of the revision (positive when they added more bytes than they removed, negative when they removed more than they added)
- **content:** the substance of the revision
- **comment:** a brief explanation/justification of the revision

We also used R to a small extent to pull daily page-views data on each article, using the *WikipediR* package[3].

4.1 Selecting articles of interest

The rich diversity in languages within the regions surrounding Jammu and Kashmir as well as the recent revocation of Article 370 of the Indian Constitution have made Wikipedia articles related to Jammu and Kashmir the focus of many new additions of information, as well as conflicts in ideas and sources of information between editors. All these factors make the following three articles very interesting to our research which is focused on editing patterns and conflict between editors of different versions of Wikipedia.

The Wikipedia articles chosen to be studied:

(1) *Article 370 of the Constitution of India*
(2) *Kashmir conflict*
(3) *Insurgency in Jammu and Kashmir*

Of these three articles, *Article 370 of the Constitution of India* and *Kashmir conflict* have Hindi, English and Urdu versions while *Insurgency in Jammu and Kashmir* only has Hindi and English versions.

4.2 Methodology

4.2.1 Page-views. We examined page-views time series for each article in order to validate our choice of articles, and to support later analysis of the correlation between the number of page-views and the number of edits on a given article at a given time. Information seekers visit the Wikipedia page on a topic following breaking news events that relate to the topic [23]. The more significant an event to a topic, the more views that topic will garner. A 2017 study of multilingual students' Wikipedia habits by Soler-Adillon and Freixa showed that their subjects favor the English Wikipedia when looking for information, even if English is not the language in which they are most fluent [21]. Their subjects were able to read and write in at least Catalan, Spanish, and English. We believe their finding extends to Hindu- and Urdu-speakers who also have English proficiency, and our interviews supported this.

4.2.2 Revisions Data. We also examined the sizes (in bytes) of the articles over time, which climb and drop as editors add and remove content. Using the revisions time series of each article, we can observe whether there exists a correlation between the edit spike times and volumes on the same article in English, Hindi and Urdu.

---

[1]https://www.mediawiki.org/wiki/API:Revisions
[2]https://github.com/ViralNotPrasad/cscw-wiki
[3]https://cran.r-project.org/web/packages/WikipediR/index.html

How quickly do edits spike after a significant event? To answer this question, we performed time-lagged cross-correlation analysis on the views and edits of each article. Taking the two daily time series (views and edits), we computed the correlation between the views on one day and the edits on the same day; the edits the day after; the edits two days after; and on for two weeks, for every day in the views series. Thus, we found how many days editing activity lags behind viewing activity for each article in each language.

We analyzed the similarity between sets of editors, across languages and between articles in the same language, by computing Jaccard coefficients [18], used previously by Twyman et al to examine the similarity between the sets of editors of Wikipedia articles [23]. This metric describes the fraction of the cumulative unique editors of two articles who edited *both* articles; that is, taking the intersection of the two sets of editors, divided by the symmetric difference. The Jaccard coefficient ranges from 0 to 1, 1 being the most similar (all in common) and 0 being the most diverse (none in common).

We also looked at where editors were editing from, in the cases where IP addresses were available. If a Wikipedia user has a username or choose to be anonymous, it is not possible to see their IP address without privileges that are only afforded to users with certain roles on the site, but some users

4.2.3   Wikipedia Editor Interviews. To extract unique editors from Wikipedia, the XTools web tool[4] was used[5]. Editors were extracted for the articles in our corpus and filtered to a unique set for each article. Editors of interest were identified using the following criteria for every article:

- Amount of data added/removed in bytes
- Time since last edit
- Total numbers of edits in an article

Thereafter, each editor barring anonymous editors was sent an recruitment email using Wikipedia's *EmailUser* facility [4].

4.2.4   Demographics. A total of 5 Wikipedia editors were interviewed for this study. Participants represented varied backgrounds and base locations. The participants also had varying levels of experience on Wikipedia with edits ranging from <5,000 total edits up to over 2 million total edits. Interestingly, all the editors that were interviewed were male.

Sl.No	Age	Country	Wiki	Lang.	Occupation	Edits	Editor Since
1	23	IN	Hi	Hi, En	Student	<5K	2019
2	22	IN	Hi	Marwari, Hi, En	Media	40-50K	2015
3	40	PK	Ur	Ur, En	Online News Editor	20-30K	2013
4	37	US	En	En, Spanish	Technical Writer	2,000,000+	2004
5	19	IN	En	En, Hi, Bengali	Student	10-20K	2010

Table 1. Editor demographics (Hi=Hindi, Ur=Urdu, En=English)

Editors were asked to participate in an hour long semi-structured interview [8]. The interviews were conducted via online call services such as Skype, Google Hangouts and Whatsapp. Three researchers joined in on the call, with one leading the interview and the other two taking notes and asking follow up questions.

---

[4]https://xtools.wmflabs.org/

The purpose of the interview was to extract editor motivations, goals and frustrations in the editing process as well their personal views on the Kashmir Conflict.Editors were also questioned on their relationships with other editors in the Wikipedia community. Using the semi-structured interview sessions, we hope to understand motivations, collaborative structures, sources of conflict and sources of references as well as editing processes of these editors in each of the English, Hindi and Urdu Wikipedias.

After the interview was completed, the researchers extracted key elements ideas discussed in the interviews using thematic analysis[10]. To help organize the interview data, key elements and ideas were coded into important themes/keywords.

We hope that these interviews will help us gain more insight into efforts being taken to improve the quality and contributions of editors on the local language versions of Wikipedia in India, which include Urdu and Hindi.
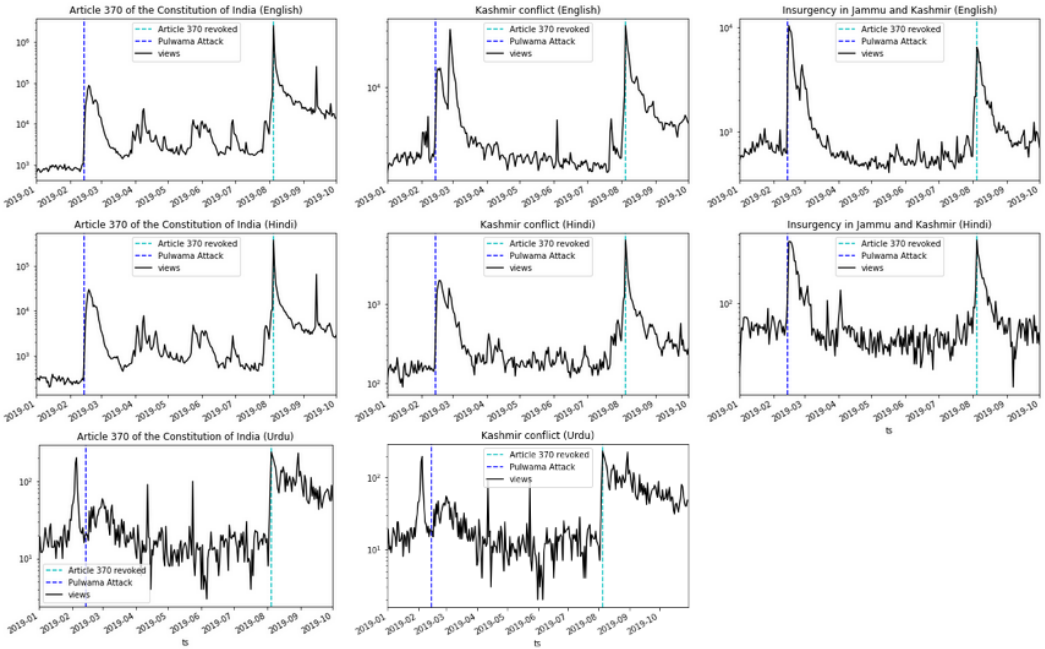

Fig. 1. Page-views per day, by article and language, log scales

5   RESULTS

5.1   Page Views:

As figure 1 shows, there were local maxima on February 14 2019 and August 5 2019 for all the articles in our corpus, with the possible exception of the Urdu articles, neither of which saw as clear of a spike in views when the Pulwama Attack occurred as did the articles in Hindi and English. There was a spike on February 5 2019, which can be attributed to Kashmir Solidarity Day, a national holiday in Pakistan [7], and both *Article 370* and *Kashmir Conflict* in Urdu did have a smaller swell in page-views a few days after the Pulwama Attack. The fact that people seeking information in Urdu on Kashmir Solidarity Day referred to the two Urdu articles in our corpus, though we were not looking for this connection, further supports that the articles we chose are of particular relevance to current events in J&K.

The magnitude of page views in 2019 is worth noting. The page views on the English *Article 370* peaked at over 1,000,000 when Article 370 was revoked. The Hindi version of the same article garnered over 100,000. By contrast, neither Urdu article exceeded 400 views. We cannot deduce that only so many Urdu- and Hindi-fluent internet users look to Wikipedia for information on these events, since it should go without saying that many Urdu- and Hindi-speakers are also proficient or fluent in English, and many will only refer to the English Wikipedia [21].

5.2 Time Series Analysis:

Figure 3 shows the spikes and volumes of edits for the Hindi version of the *Article 370 of the Constitution of India*. The time frame captured in the figure is a very good example of the correlation between a real-time event relevant to the page (Revocation of the Article 370 by the Government of India on the 5th of August 2019), and an edit spike involving several edits (mostly additions) explaining this move by the Indian Government.
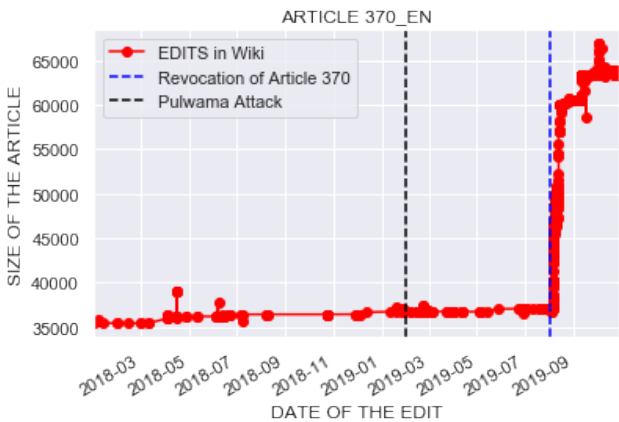

Fig. 2. Edit spikes and volumes for the Article 370 of the Constitution of India - English Version
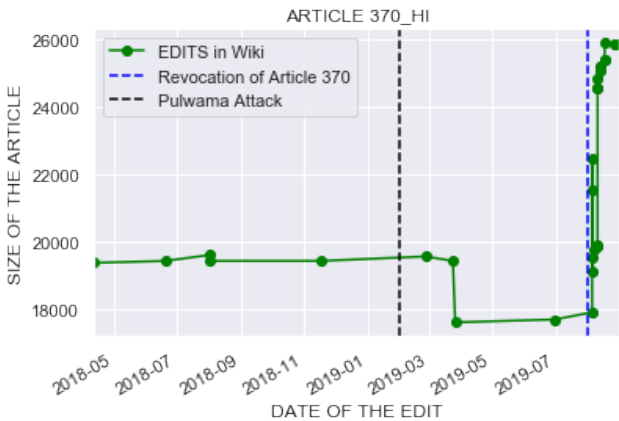

Fig. 3. Edit spikes and volumes for the Article 370 of the Constitution of India - Hindi Version

Following however, figure 2 seems to paint the complete picture showing the edit spikes around August 2019 as well with a huge number of edit spikes concerning the same event. This goes to address RQ1(a) that spikes over time in editing volume will be correlated in all three languages for the same article.

5.3 Response Time

The pages on *Article 370* were the only ones that showed significant correlation between the volume of edits and page-views over time. The peak correlations indicated in figure 4 are all significant at p<0.001. Offset refers to how many days the time series were offset by, e.g. the views and edits series for the Hindi version of the article are most synchronous when the views series leads by five days.

On the English article, edits and views are most highly correlated on the same day (r=0.86); that is, when page-views are high on a give day, edits, too, are high on that day (and the inverse is true as well). Editors of the English *Article 370* respond quickly to increased interest in a page, which as we saw, tends to signify an important event has just occurred. Edits of the Hindi version are also highly correlated (r=0.55) with same-day views, although the tallest peak correlation occurs with a five-day lag (r=0.60), meaning Hindi daily editing volume tends to spike five days after page-views spike. The Urdu version of the article sees less correlation between page-views and edits (r=0.31 with a two-day lag, r=0.25 with a three-day lag), but still significant at p<0.001.

English is the lingua franca—it is not surprising that on an English article on an important topic like Article 370, editors respond quickly to new events. It is surprising that despite having so many fewer editors, the Hindi and Urdu pages are also edited shortly after page-views spike.
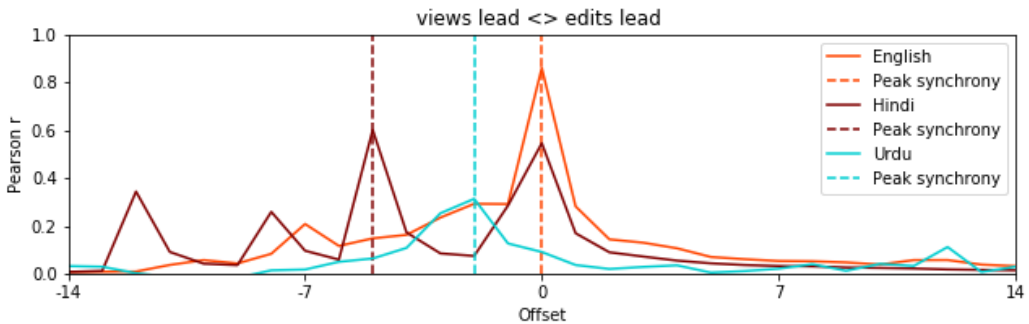

Fig. 4. Time-lagged cross-correlation between edits and page-views on Article 370, in all three languages

5.4 Editor Similarity:

To address RQ2, we examined the overlap between sets of editors of the two articles in our corpus that existed in all three languages. We found that there were few editors who made revisions in two languages, and none who made revisions in both Hindi and Urdu. The Urdu articles on *Kashmir Conflict* and *Article 370* have been edited by under twenty editors each (five of whom they share), since their respective inceptions in 2013 and 2018. The Hindi article on *Article 370* began in 2011, three years earlier than the Hindi article on *Kashmir Conflict*, although the latter has had more editors. The English articles have existed longest and seen an order of magnitude more editors than Urdu and Hindi combined. This supports our H5: the number of editors increases from Urdu to Hindi to English.

Table 3 shows the similarity between editor sets for each article, by language. For example, 0.03 indicates that merely three in a hundred editors of the English *Article 370* page also edited either the English *Kashmir Conflict* page or the *Insurgency* page.

We expected to find that a user who revised an article in either Hindi or Urdu would be likely to have also edited the corresponding article in English. However, contrary to our H4, we found that between languages, the articles shared almost no editors. The Jaccard coefficients in these cases were minuscule (on the order of .001); instead, table 4 presents the actual numbers of editors in common.

Language	Article	Editors	Page creation date
Hindi	*Article 370*	42	2011-03-12
Urdu	*Article 370*	10	2018-09-27
English	*Article 370*	503	2006-02-13
Hindi	*Kashmir Conflict*	51	2014-01-31
Urdu	*Kashmir Conflict*	17	2013-10-18
English	*Kashmir Conflict*	1,577	2005-10-29
Hindi	*Insurgency*	33	2011-04-14
English	Insurgency	641	2005-05-20
Urdu	*Insurgency*	NA	NA

Table 2. Number of unique editors and page creation date of each article

	Jaccard coefficients		
Language	Article 370	Kashmir Conflict	Insurgency
Hindi	0.15	0.15	0.13
Urdu	0.4*	1*	NA
English	0.03	0.23	0.10

Table 3. Similarity within-language between the editors of the *Kashmir Conflict*, *Article 370*, and *Insurgency in Jammu and Kashmir* pages

6 DISCUSSION

6.1 Interview Results

A thematic analysis of the interviews transcripts as well as notes and observations from the interviews yielded the following key themes: *Motivations; Collaboration Strategies; Sourcing Information, Structure, Content & Consistency (Between Languages); and Conflict, Propaganda & Vandalism*

Between	Kashmir Conflict	Article 370	Insurgency
Hindi and English	1	3	3
English and Urdu	2	0	NA
Urdu and Hindi	0	0	NA

Table 4. Editors in common across languages, by article

6.1.1 Motivations. All five interviewed Wikipedia editors indicated that editing Wikipedia articles was a hobby and a passion. Beyond this individual editors had additional reasons for actively engaging in the editing process. Participants from the Hindi and Urdu versions of Wikipedia were interested in "Growing the language" and "Providing a service to readers of their language." One participant indicated that he was interested in making knowledge more accessible; furthermore, the participant indicated that privacy and anonymity along with freedom of access to free knowledge were additional motivations for editing Wikipedia. In general, participants were interested in improving the quality and size of Wikipedia articles.

6.1.2 Collaboration Strategies. According to Sundin [22], gathering, discussing and disputing knowledge claims are important parts of editor activities on Wikipedia. These knowledge claims are collaboratively constructed through situated practices of many actors.

When it came to collaboration, all the Hindi and Urdu Wikipedia stressed the importance of communicating with their peers. This was not necessarily true for the two English Wikipedia editors interviewed, who preferred working alone on their edits.

The most visible collaborative interactions happened on the article talk page or on the UserTalk pages on Wikipedia. These pages act as a forum for communications about articles or bilateral communication between editors, though the purpose of conversations happening on these pages was mostly record keeping. The Hindi and Urdu editors used various virtual and offline channels to communicate with one another. Virtually, Hindi and Urdu editors preferred Facebook and Whatsapp for communication. These channels were usually used for informal communication, delegating editing work, and discussing issues on Wikipedia editing. The interviews also revealed that many Hindi and Urdu editors meet up regularly, especially if they are living in close proximity. The English editors interviewed were not actively part of any offline or virtual groups related to Wikipedia, although one participant mentioned the success of Wikipedia collaborations in some cases where external organisations were involved such as Museums.

Although many Wikipedia conferences and groups exist, our participants were not active members of these organisations, or did not see these forums as their primary method of collaboration.

It was very interesting to note that the Hindi and Urdu editors interviewed felt a stronger sense of community through offline and virtual collaboration. We believe that the reasons for this strong collaboration are the common goals of generating information on a rather small version of Wikipedia and the small number of editors that exist on the platform. While the English Wikipedia has many groups as well, there seems to be a different focus on the editing process, with editors being more focused on topics that they find most interesting or that they know most about.

Additionally, in the Hindi and Urdu versions of Wikipedia, a common thread of observing and guiding new users was found. These editors were more invested in the addition of newer editors on the platform when compared to the English editors.

Cultural differences between the editors may have a role to play in the way they focus on generating Wikipedia data and collaboration. With most Hindi and Urdu editors located in culturally similar regions, it is likely that cultural practices have an effect on collaborative practices as well. This idea of a digital natives (persons brought up in the age of computer technology) in the global south differing from those in the global north has been echoed in previous research [13].

6.1.3 Sourcing Information. Before interviewing our participants, we expected to see some sort of bias or differences in references chosen between editors of different languages on the same topic. Instead, we found something else entirely. All the Hindi and Urdu Wikipedia editors interviewed indicated that their primary source of references was from the English Wikipedia, with some content being directly translated from the English version of the Wikipedia article. Sometimes, if the editors were familiar with a topic or if they personally knew that a local source was more correct

than the one in the English Wikipedia, they would use that reference instead. But more often than not, editors were focused on sourcing information directly from the English Wikipedia articles.

On the other hand, English editors usually sourced information from "reliable sources" such as reputed news organisations, public databases and other relevant sources such as journals and other scholarly articles. The reliability of sources was determined through various online forums and agreed upon standards. Conversely, it may be that local sources may not necessarily be seen as "reliable sources" in the fact that they may lie outside the linguistic and social access of most editors, an example may be small newspapers from countries in the global south such as India and Pakistan [13]. Wikipedia editors must understand, not just how to describe a subject, but how to argue that the new material deserves to be added.[13]

This finding was one of the more interesting from our interviews. It indicates that Hindi and Urdu editors are interested in adding to their versions of Wikipedia and increasing the accessibility of knowledge to native Hindi and Urdu speakers. This corresponds to our previous findings of the motivations of Wikipedia editors.

6.1.4 Structure, Content & Consistency. The English editors interviewed indicated that their focus on the editing process had moved more towards maintaining the structure of the content, fixing the writing style of articles, and adding information when necessary, whereas the Hindi and Urdu editors focused more on adding content, and therefore their pages are very relaxed in building rules and structure for their versions of Wikipedia. The idea was that the pages don't have to pretty; they just have to be there.

There are, however, house rules that are build and maintained by admins from each versions of Wikipedia. One of the participants indicated that his admin role had allowed him to build house rules that were inspired from the English Wikipedia but with additions more relevant to their version of Wikipedia.

We believe that as Wikipedia versions in their nascent stages, the Hindi and Urdu are not yet at the same level of maturity in terms of structure and formatting; therefore, one can expect a certain lack of observable quality in their pages at present, but the quality will increase as the Wikipedia versions mature.

6.1.5 Conflict, Propaganda Vandalism. Conflicts between editors usually happen because of differences in opinions on content, references, and templates, but also, in many cases, because of vandalism and propaganda. Wikipedia has a neutral point of view policy on its articles, to which most active editors try to adhere [2], although many vandals and propagandists ignore this rule.

It was found from anecdotal evidence that intralanguage conflicts for the Hindi and Urdu versions of Wikipedia were largely resolved through offline and virtual communications, while on-the-record (on *UserTalk* and *article Talk* pages) conflicts were less common than on the English Wikipedia.

The English Wikipedia editors gave anecdotal evidence of many differences in opinion on the structure, semantics and references, although none lasted very long. These disputes are largely won through a war of attrition, in that editors who stand their ground on the *talk* pages for longer tend to get their way.

All our participants agreed that vandalism was a major issue plaguing Wikipedia, although any one incidence of vandalism will not persist for long, as active editors ensure that mistakes and misinformation are removed. Editors from all three Wikipedias gave anecdotal evidence of vandalism and propaganda propagating through editors without usernames, i.e., who use I.P. addresses to make edits.

One English Wikipedia editor suspected that there might be various editors using I.P. addresses spreading propaganda on behalf of governments. Another English Wikipedia editor confirmed this

by giving anecdotal evidence of a group of rogue editors (using I.P. addresses) spreading propaganda on the day of the revocation of article 370. Further inquisition into the topic revealed that these rogue editors were likely coordinating their edits as there were similar edits on multiple pages at the same time.

The Hindi and Urdu Wikipedia editors also complained of propaganda based edits through I.P. addresses although this was not as prevalent. The Urdu editor also mentioned that a lot of political propaganda take place on controversial articles in the Urdu version, although these instances are removed by active editors.

In summary, vandalism and propaganda, while prevalent in Wikipedia, are rarely added by active Wikipedia users. On the other hand there is reason to believe that there may be coordinated efforts to spread misinformation and propaganda through I.P. address accounts.

7   CONCLUSION

In all three languages, when traffic to a given page increases in the wake of an event significant to the article, editors respond within a few days to with updates to the article, although at different rates. The editors we interviewed were unilaterally devoted to the NPOV that is Wikipedia's trademark. Vandalism, though prevalent according to the editors we interviewed, is swiftly undone. Collaboration differs significantly between editors of English Wikipedia articles from the editors of Hindi and Urdu Wikipedia articles, where editors show more of a sense of community. To summarize, Wikipedia editors in all three languages (English, Hindi and Urdu) consistently work together on maintaining accurate, neutral, and up-to-date information that is becoming more and more accessible to readers around the world.

8   FUTURE WORK

We plan to extend the scope of our work by applying our methods to more articles, as well as interviewing more editors.

A closer look at edits by anonymous users would complement our work. Interviewees told us that most vandalism is committed by editors without usernames (only IP addresses), and several speculated that some of these editors might be paid propagandists, which is a growing threat, not only to Wikipedia and not only in the region we chose for this study, but to websites and organizations around the globe.

Another direction to explore would be to study, broadly, how new language versions of Wikipedia are developed; how "house rules" are agreed on (or not); and why some projects fail (due to propagandists outnumbering regular editors, or other factors), as the Bulgarian Wikinews project recently did[3].

ACKNOWLEDGMENTS

REFERENCES

[1] [n.d.]. Kashmir shuts down on Burhan Wani's death anniversary; Yatra, security convoy movement suspended - The Economic Times. https://economictimes.indiatimes.com/news/politics-and-nation/kashmir-shuts-down-on-burhan-wanis-death-anniversary-yatra-security-convoy-movement-suspended/articleshow/70131204.cms?from=mdr. (Accessed on 12/12/2019).
[2] [n.d.]. Neutral Point of View. https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view. (Accessed on 12/12/2019).
[3] [n.d.]. Proposals for closing projects/Deletion of Bulgarian Wikinews. https://meta.wikimedia.org/wiki/Proposals_for_closing_projects/Deletion_of_Bulgarian_Wikinews (Accessed on 12/13/2019).

[4] [n.d.]. Wikipedia:Emailing users - Wikipedia. https://en.wikipedia.org/wiki/Wikipedia:Emailing_users. (Accessed on 12/12/2019).

[5] [n.d.]. XTools. https://xtools.wmflabs.org/. (Accessed on 12/12/2019).

[6] 2019. *The Gazette of India : Extraordinary*. Government of India. http://egazette.nic.in/WriteReadData/2019/210049. pdf

[7] 2019. Kashmir Solidarity Day. https://en.m.wikipedia.org/wiki/Kashmir_Solidarity_Day

[8] K Louise Barriball and Alison While. 1994. Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing-Institutional Subscription* 19, 2 (1994), 328–335.

[9] bbc com. 2019. Pulwama attack: India will 'completely isolate' Pakistan. Retrieved February 28, 2008 from https://www.bbc.com/news/world-asia-india-47249133

[10] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. (2012).

[11] dawn.com. 2019. On Kashmir attack, Shah Mahmood Qureshi says 'violence is not the govt's policy. Retrieved December 2, 2019 from https://www.dawn.com/news/1464205

[12] Nuša Farič and Henry WW Potts. 2014. Motivations for Contributing to Health-Related Articles on Wikipedia: An Interview Study. *J Med Internet Res* 16, 12 (03 Dec 2014), e260. https://doi.org/10.2196/jmir.3569

[13] Heather Ford and R Stuart Geiger. 2012. Writing up rather than writing down: Becoming Wikipedia literate. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. ACM, 16.

[14] Scott A. Hale. 2014. Multilinguals and Wikipedia Editing. In *Proceedings of the 2014 ACM Conference on Web Science (WebSci '14)*. ACM, New York, NY, USA, 99–108. https://doi.org/10.1145/2615569.2615684

[15] Brent Hecht and Darren Gergle. 2009. Measuring Self-focus Bias in Community-maintained Knowledge Repositories. In *Proceedings of the Fourth International Conference on Communities and Technologies (C&T '09)*. ACM, New York, NY, USA, 11–20. https://doi.org/10.1145/1556460.1556463

[16] Brent Hecht and Darren Gergle. 2010. The Tower of Babel Meets Web 2.0: User-generated Content and Its Applications in a Multilingual Context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 291–300. https://doi.org/10.1145/1753326.1753370

[17] Joseph J Hobbs. 2008. *World regional geography*. Nelson Education.

[18] Paul Jaccard. 1912. The Distribution Of The Flora In The Alpine Zone.1. *New Phytologist* 11, 2 (Feb 1912), 37–50. https://doi.org/10.1111/j.1469-8137.1912.tb05611.x

[19] Brian Keegan. 2014. *Emergent Social Roles in Wikipedia's Breaking News Collaborations*. Springer.

[20] Ulrike Pfeil, Panayiotis Zaphiris, and Chee Siang Ang. 2006. Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication* 12, 1 (2006), 88–113.

[21] Pere Soler-Adillon, Joan; Freixa. 2017. Wikipedia access and contribution: Language choice in multilingual communities. A case study. *Anàlisi* 57 (2017).

[22] Olof Sundin. 2011. Janitors of knowledge: constructing knowledge in the everyday life of Wikipedia editors. *Journal of documentation* 67, 5 (2011), 840–862.

[23] Brian C.; Shaw Aaron Twyman, Marlon; Keegan. 2016. Black Lives Matter in Wikipedia: Collaboration and Collective Memory around Online Social Movements. (2016).

[24] K Venkataraman. 2019. Explained | How the status of Jammu and Kashmir is being changed. *The Hindu* (Aug 2019). https://web.archive.org/web/20190808005453/https://www.thehindu.com/news/national/other-states/explained-how-the-status-of-jammu-and-kashmir-is-being-changed/article28822866.ece?homepage=true