

The Gender Gap in Wikipedia Talk Pages*

Benjamin Cabrera, Björn Ross, Marielle Dado, Maritta Heisel

Department of Computer Science and Applied Cognitive Science
University of Duisburg-Essen
firstname.lastname@uni-due.de

Abstract

Wikipedia is an important source of information in today's world. Yet, the lack of gender diversity in its community has been shown to affect the topics covered. Each Wikipedia article has a talk page that volunteer editors use to discuss proposed changes. Research on the gender bias has focused on article contribution and topic coverage, but not talk page activity. It has been suggested that the conflicts that take place in talk pages are especially intimidating for women, but this assertion has not been quantified yet. To fill this gap, we collected a dataset of all comments on Wikipedia talk pages, enriching it with gender information available from users who have chosen to disclose their gender on their user profiles or settings. Among the users active in talk pages, 49,387 indicated that they are male while only 5,996 indicated that they are female. The comments of these users make up for 4 million comments, approximately one quarter of all comments on Wikipedia. In addition, we observed that female participation varies by topic, reflecting traditional gender stereotypes: compared to Science, Technology, Engineering and Mathematics (STEM) topics, women were more active in categories such as Gender studies or Feminism. Results also indicate that a post on a talk page is 2.4% less likely to be replied to if the author is female. Likewise, reply probability varies from topic to topic. These results provide quantitative support for a gender bias in Wikipedia talk pages, and serve as a basis for discussing why overall female participation is low.

Introduction and Related Work

Wikipedia is a free online encyclopedia that is curated by millions of contributors worldwide. Given Wikipedia's vision¹ to become a platform for anyone to "freely share in the sum of all knowledge", it is reasonable to expect the Wikipedia editors to be as diverse as the general population. However, an overwhelming majority of Wikipedia editors are male, with surveys placing the number of female Wikipedia editors at only 12 to 16% (Hill and Shaw 2013). Consequently, Wikipedia has been criticized for having a gender bias that further introduces biases present in the real

world. For instance, compared to biographies of men, biographies of women on Wikipedia are more likely to be missing (Reagle and Rhue 2011). Although globally notable women are now being represented in Wikipedia, there seem to be more biographies of men that are notable in smaller, local communities, implying a glass ceiling that women must surpass in order to be included (Wagner et al. 2016). Furthermore, women are also more often characterized around personal information such as childbirth, family, marriage and divorce (Bamman and Smith 2014; Wagner et al. 2015). Women also tend to contribute less to topics that are traditionally male-dominated, such as science and engineering (Lam et al. 2011). One reason for the gender gap could be the fear of criticism or being in conflict with others, a concern which makes it less likely for some female editors to contribute (Collier and Bear 2012). Contributions by female newcomers are more likely to be reverted and women are more likely to be indefinitely blocked, hinting at a "culture resistant to female participation" on Wikipedia (Lam et al. 2011). Female editors were also less inclined to contribute if they believed no other women were contributing as well (Shane-Simpson and Gillespie-Lynch 2017). Thus, the low participation of women in Wikipedia can be attributed to a perceived unwelcoming environment towards women, which results in less coverage of women-centric topics as and the characterization of women in stereotypical ways.

Existing research (Lam et al. 2011; Wagner et al. 2015) on the causes of Wikipedia's gender gap has mostly focused on the articles themselves (e.g., edit history). So far, no studies have explicitly investigated the gender gap in the talk pages that accompany each Wikipedia article. According to the community guidelines, editors are encouraged to use the talk pages when they want to dispute a contribution in an attempt to reach consensus². Some studies have observed differences in the way men and women express themselves in talk page posts. Female editors tend to engage in more positive, emotional and relationship-oriented speech than men (Laniado et al. 2012; Iosub et al. 2014). However, these studies on gender differences in Wikipedia talk pages only included the most active editors in their data, namely those

*This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant No. GRK 2167, Research Training Group "User-Centred Social Media".

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹<https://wikimediafoundation.org/wiki/Vision>

²https://en.wikipedia.org/wiki/Wikipedia:BOLD,_revert,_discuss_cycle

who have written more than 100 comments. This is a threshold that relatively few female editors have reached (Laniado et al. 2012), which means that the sample probably does not include the large number of female editors that perceive Wikipedia to be an intimidating environment (Collier and Bear 2012). Since talk pages are meant as a forum to discuss potentially controversial article contributions, we investigate the gender gap in Wikipedia talk pages by considering the relationship between an editor’s gender and the replies that he/she receives. Furthermore, our analysis is based on the data of all Wikipedia editors that have posted on the talk page, instead of only those who are active and highly motivated to participate.

This article therefore addresses the following research questions:

RQ1: Is there a difference between the number of male and female article editors in Wikipedia talk pages, and are there differences per category?

RQ2: Is there a difference in the probability of editors receiving a reply on their posts based on their gender?

RQ3: Is the gender of a post’s author related to the gender of those who reply to it?

Next, we discuss our dataset and the method for collecting gender information and extracting talk pages. This is followed by the results of our analysis and a discussion of the implications of our results on the gender gap in Wikipedia talk pages. The code and dataset are available on Github³.

Dataset

To address the research questions, three datasets were collected and combined for the analysis: all comments made by Wikipedia editors in talk pages, the gender of these editors in case they disclosed it, and articles belonging to several categories.

A Wikipedia talk page is simply a separate page to be edited by users starting new discussion threads or replying to existing comments. Its content is not stored in a structured format, e.g. a relational database, so the page first has to be parsed. To reliably do so in an automated fashion, the GraWiTas⁴ (Cabrera, Steinert, and Ross 2017) parser tool was used. In contrast to previous research on Wikipedia talk pages, this procedure allowed us to parse an XML dump of the entire English Wikipedia as of 1 December 2017. In total, 16,974,444 comments made by 1,598,796 users on 1,841,417 articles were extracted. For each comment, its author, timestamp, article and text were extracted and stored in a relational database.

We used 2 sources to identify the gender of Wikipedia editors. First, Wikipedia allows users to store their gender to personalize their user interface. Users who choose to do so are informed that this data will be publicly available. The user preference can be accessed via a public API⁵. Please note that this preference setting, our main source of

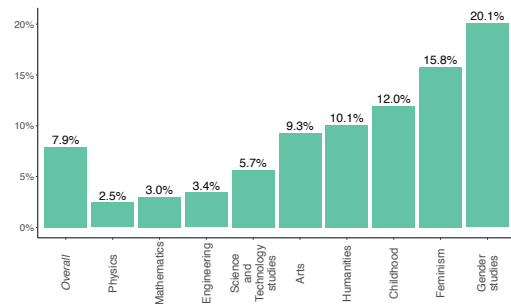


Figure 1: Proportion of comments by women on different categories

gender information, assumes gender to be binary, with the only other choice being “unknown”, the default. Secondly, Wikipedia users can create personal user pages and place “user boxes” on them, which highlight a particular characteristic of the user. Several user boxes deal with gender or allow inferring the gender of the user. Crawling the API yielded gender information for 50,829 users (45,406 male, 5,423 female), while gender information for 11,696 users (9,983 male, 1,713 female) was collected from user boxes. Removing duplicates, the genders of 55,383 talk page users (49,387 male, 5,996 female) were identified. Thus, about 3% of the users who commented in talk pages disclosed their gender, but these users make up for 4,047,842 comments, i.e. 23.8% of all comments on Wikipedia.

To analyse differences between the genders with regard to different topics, we used the Wikipedia category system. Articles can be linked to special category pages. The categories themselves can be members of other categories or contain links to subcategories. Using the category API⁶, all articles of a given category were extracted as well as recursively for subcategories. This recursive process was stopped at depth 2 to avoid issues stemming from the fact that in general, the Wikipedia category→subcategory relationship does not form a tree but a cyclic graph.

Results

Talk Page Participation by Gender and Topic

RQ1 asks if and how the gender gap in user and edit numbers translates to the talk pages. Of the 4,047,842 comments for which the gender of the author could be identified, 7.9% were by women. Overall, women make up 10.8% of all users that have at least commented once on a talk page. This shows that women indeed write fewer comments, not only in absolute terms but also when normalizing for the user gender gap. Comments also vary in length with respect to the author’s gender. Comments by female editors contain a mean of 529 characters (SD = 804.0, median = 318), and comments by male editors only 464 (SD = 675.9, median = 285).

RQ1 also asks whether there are any differences in the gender gap per topic. Figure 1 shows the proportions of comments made by men and women in different categories.

³<https://github.com/bencabrera/icwsm2018>

⁴<https://github.com/bencabrera/grawitas>

⁵<https://www.mediawiki.org/wiki/API:Users>

⁶<https://www.mediawiki.org/wiki/API:Categories>

| Top-level posts by ... | Number of posts |
|-------------------------------------|---------------------|
| Male authors | 1,641,282 |
| — which received at least one reply | 574,980 (35.03%) |
| Female authors | 143,815 |
| — which received at least one reply | 49,163 (34.18%) |

Table 1: Reply probability by gender

Comments across topics are male dominated in absolute terms. Even in categories such as Feminism, women only account for 16% of all comments. The largest gender gap of the tested categories can be found in STEM categories such as Physics, Engineering or Science and Technology studies.

Reply Probability by Gender

The substantial difference between the numbers of male and female editors has been hypothesized to be due to a climate on talk pages that is hostile towards women (Lam et al. 2011). One way in which such a bias against female editors might become apparent is in reply probability (RQ2). One would expect the chance of getting a reply to a comment to be independent of the gender of the original author of the comment.

To avoid distorting the results, only top-level posts were considered in this analysis. Top-level posts are those that begin a new discussion thread, in contrast to those made in reply to an existing post. If the latter does not receive a reply, this could simply indicate that the discussion has been resolved.

There is indeed a difference between top-level posts by male and female editors in their probability of receiving a reply (see Table 1). If the author of a post is female, it was 2.4% less likely to receive at least one reply in our dataset than if the author was male ($0.3418/0.3503 = 0.976$)⁷. A logistic regression with gender as the only independent variable reveals that its effect on reply probability is statistically significant ($z = -6.462$, $p = 1.03 \times 10^{-10}$, $\beta = -0.037$, $\exp(\beta) = 0.963$).

Figure 2 shows the ratio between the reply probabilities for female and male editors by category. For example, in the engineering category, posts on talk pages by female editors (anywhere in the discussion thread) were 25.6% less likely to be replied to than those by men. The differences are significant at a Bonferroni-adjusted α level of 0.05/10 for the categories Engineering ($z = -3.746$, $p = 1.8 \times 10^{-4}$, $\beta = -0.435$, $\exp(\beta) = 0.648$) and Childhood ($z = -3.623$, $p = 2.91 \times 10^{-4}$, $\beta = -0.179$, $\exp(\beta) = 0.837$).

Gender of the Author of a Reply

Examining the gender of the editor replying to a post (RQ3), we found that male editors are more likely to reply to other

⁷To avoid ambiguity, the reported numbers, unless otherwise stated, are the ratios between two probabilities also known as risk ratios, not absolute differences between probabilities or odds ratios.

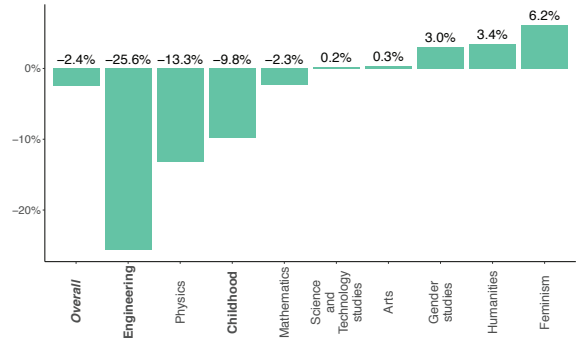


Figure 2: Ratio of reply probabilities of top-level posts by female and male editors, by category. For example, overall, women’s comments are 2.4% less likely to be replied to than men’s. Bold category names mark statistically significant differences ($p < 0.005$).

| Gender of reply (→) and orig. comment (↓) | Male | Female |
|--|-----------------|---------------|
| Male | 488,585 (92.6%) | 39,280 (7.4%) |
| Female | 39,183 (87.7%) | 5,514 (12.3%) |

Table 2: Number of replies distinguished by the authors’ genders. Comment/reply pairs were only considered if the gender was known for both authors. Replies to one’s own comments were excluded.

male editors while female editors are more likely to reply to other female editors. The results are presented in Table 2. For example, if the author of a post is female, the probability that the reply was written by a another female editor is 12.3%. However, if the original poster is male, the probability of the author of the reply being female is only 7.4%.

Discussion and Outlook

In this paper, we investigated the gender gap in Wikipedia by analyzing the talk pages, specifically the probability of receiving replies depending on the author’s gender.

Addressing RQ1, women are clearly outnumbered on talk pages by a factor of 10 to 1. Women also write fewer comments, in relation to the proportion of women that have at least commented once on a talk page. However, they tend to be longer than comments by men. Regarding RQ2, we found that women are less likely to receive replies in talk pages than men. While a difference in the reply probabilities of 2.4% might not be large enough to be noticed by an individual editor, the number of affected comments in an online encyclopedia with millions of comments is in the order of tens of thousands. In addition, the differences in reply probability were much larger for some topics. Male editors are more likely to be replied to when talking about engineering, a traditionally male-dominated topic for which the number of female editors on Wikipedia is also much lower than average. Perhaps surprisingly, for childhood, a stereotypically female-dominated topic with a larger-than-average number

of female editors (see Figure 1), posts by men were also more likely to be replied to. Finally, men are more likely to reply to men, and women to women (RQ3). This observation confirms the earlier result that the talk page network is assortative with respect to gender (Laniado et al. 2012), but on a much larger dataset.

Together, these results suggest one possible cause for the gender gap in talk pages and perhaps even Wikipedia at large. The talk page is one of the more social aspects of Wikipedia, as editors have an opportunity to directly communicate with each other. Groups often exhibit homophily with respect to attributes such as gender (McPherson, Smith-Lovin, and Cook 2001). However, this finding has important ramifications for Wikipedia. If members of online communities are naturally inclined to socialize with other individuals of the same gender, then in a community that is predominantly male such as Wikipedia, prospective female members will find it especially hard to participate. This tendency has been seen when it comes to article contributions: women are less likely to contribute if they believe no other women are present (Shane-Simpson and Gillespie-Lynch 2017). Since new high-volume Wikipedia editors place special importance on social motives and reputation (Balestra et al. 2016), a low reply rate could contribute to turning away some of the potentially most productive members of the community. As a consequence, Wikipedia editors might need to make more of an effort to reply to posts on talk pages regardless of the author's gender, especially in STEM-related topics such as engineering.

The present work is, to our knowledge, the only one to date that has addressed the gender gap on Wikipedia talk pages using a dataset covering the entire English Wikipedia instead of a sample of articles and/or editors. Of course, other language editions might exhibit different gender bias. Although our dataset is larger than in other studies on Wikipedia talk pages, our analysis is limited to the editors whose gender can be determined (about a quarter of comments). A recent study suggests that editors take cues on fellow editors' gender based on their usernames (Shane-Simpson and Gillespie-Lynch 2017). In ongoing work we are therefore also exploring *perceived* gender and how this influences the rate of talk page replies.

This article thus contributes to the ongoing discussion on reasons why female participation is low and suggests possible reasons. To further investigate these issues, the data will have to be monitored over time and the contents of the posts will have to be analyzed. The same data will also make it possible to observe whether the gender gap is shrinking. Finally, we hope that this paper will further raise awareness of this issue and contribute to the discussion of how it can be addressed.

References

- Balestra, M.; Arazy, O.; Cheshire, C.; and Nov, O. 2016. Motivational Determinants of Participation Trajectories in Wikipedia. In *Proceedings of ICWSM*, 535–538. Palo Alto, CA: The AAAI Press.
- Bamman, D., and Smith, N. A. 2014. Unsupervised discovery of biographical structure from text. *Transactions of the Association for Computational Linguistics* 2:363–376.
- Cabrera, B.; Steinert, L.; and Ross, B. 2017. GraWiTas: a Grammar-based Wikipedia Talk Page Parser. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017, Valencia, Spain, April 3-7, 2017, Software Demonstrations*, 21–24. Stroudsburg, PA: ACL.
- Collier, B., and Bear, J. 2012. Conflict, criticism, or confidence: An empirical examination of the gender gap in wikipedia contributions. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW '12*, 383–392. New York, NY, USA: ACM.
- Hill, B. M., and Shaw, A. 2013. The Wikipedia gender gap revisited: Characterizing survey response bias with propensity score estimation. *PLOS ONE* 8(6):1–5.
- Iosub, D.; Laniado, D.; Castillo, C.; Fuster Morell, M.; and Kaltenbrunner, A. 2014. Emotions under discussion: Gender, status and communication in online collaboration. *PLOS ONE* 9(8):1–23.
- Lam, S. T. K.; Uduwage, A.; Dong, Z.; Sen, S.; Musicant, D. R.; Terveen, L.; and Riedl, J. 2011. WP:Clubhouse?: An exploration of Wikipedia's gender imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration, WikiSym '11*, 1–10. New York, NY, USA: ACM.
- Laniado, D.; Kaltenbrunner, A.; Castillo, C.; and Morell, M. F. 2012. Emotions and dialogue in a peer-production community: the case of wikipedia. In *Proceedings of the 8th Annual International Symposium on Wikis and Open Collaboration, WikiSym '12*. New York, NY, USA: ACM.
- McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27(1):415–444.
- Reagle, J., and Rhue, L. 2011. Gender bias in Wikipedia and Britannica. *International Journal of Communication* 5:21.
- Shane-Simpson, C., and Gillespie-Lynch, K. 2017. Examining potential mechanisms underlying the Wikipedia gender gap through a collaborative editing task. *Computers in Human Behavior* 66:312–328.
- Wagner, C.; Garcia, D.; Jadidi, M.; and Strohmaier, M. 2015. It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. In *Proceedings of ICWSM*, 454–463. Palo Alto, CA: The AAAI Press.
- Wagner, C.; Graells-Garrido, E.; Garcia, D.; and Menczer, F. 2016. Women through the glass ceiling: gender asymmetries in wikipedia. *EPJ Data Science* 5(1):5.