Gender Differences in Wikipedia Editing

Judd Antin[†]
Yahoo! Research
Santa Clara, VA 95054
jantin@yahoo-inc.com

Raymond Yee
Data Unbound LLC
Pittsburgh, PA 15217
raymond.yee@gmail.com

Coye Cheshire
UC Berkeley iSchool
Berkeley, CA 94720
coye@ischool.berkeley.edu

Oded Nov
Polytechnic Institute of New
York University
Brooklyn, NY 11201
onov@poly.edu

ABSTRACT

As Wikipedia has become an indispensable source of online information, concerns about who writes, edits, and maintains it have come to the forefront. In particular, the 2010 UNU-MERIT survey found evidence of a significant gender skew: fewer than 13% of Wikipedia contributors are women. However, the number of contributors is just one way to examine gender differences in contribution. In this paper we take a more fine-grained perspective by examining how much and what types of Wiki-work men and women tend to do. First, we find that the so-called "Gender Gap" in number of editors may not be as wide as prior studies have suggested. Second, although more than 80% of editors in our sample were men, among the bottom 75% of editors by activity-level, we find that men and women made similar numbers of revisions. However, among the most active Wikipedians men tended to make many more revisions than women. Finally, we find that the most active women in our sample tended to make larger revisions than the most active men. We conclude by discussing directions for future research.

Categories and Subject Descriptors

H.5.m [**Information interfaces and presentation (e.g. HCI)**]: Miscellaneous

General Terms

Human Factors

Keywords

Wikipedia, gender, participation

1. INTRODUCTION

Since the inception of Wikipedia, many have asked the question "Who writes Wikipedia?" Over time many answers to

this question have been proposed. However, in 2010 a specific and to some troubling answer to the question of who writes Wikipedia came to the forefront: predominantly men. The 2010 UNU-MERIT study of worldwide Wikipedia users found that less than 13% of contributors were female [4]. A re-surfacing of this report in an early 2011 New York Times article [1] led to a flurry of media attention to Wikipedia's so-called "Gender Gap". The original report cited no reasons for the drastically different numbers of male and female contributors, but in ensuing discussions many expressed their own opinions and speculated about the reasons for the gap (See, e.g. [3]).

While sheer number of contributors is an important metric, it is just one way to examine gender differences on Wikipedia. In this paper we focus not on the gender gap in terms of quantity of editors but on the potential gap in terms of the quantity and substantive type(s) of Wiki-work that men and women tend to do in their first experiences with Wikipedia. We also examine this alternative gender gap across the spectrum of Wikipedia contributors — from the numerous Wikipedians who edit infrequently to the few who take on a heavy editing burden. We look for differences between men and women's editing activity in terms of the number and size of the revisions they make, as well as the substantive type of editing work represented by those revisions. Finally, we discuss the implications of our results and present directions for future research.

2. BACKGROUND

For more than 100 years scholars in a variety of disciplines have examined gendered work (See, e.g. [10]). More recently, researchers have also examined gender differences in various types of online participation (See, e.g. [6]). A variety of theories have been advanced to help explain the prevalence of gender disparities online and offline. A key insight from this body of research is that the processes by which gender disparities occur need not be conscious. For example, "occupational stereotypes" [5] — preconceptions about the nature of specific jobs and the gender, ethnicity, or other descriptive characteristics of the people who do them — have been used to help explain ongoing gender disparities in education, employment, and leisure activity. These cultural biases can influence an individual's attitudes about

implicitly or explicitly joining the group of people who do a specific type of work. Since in-group identification can promote cooperation [2], this gender stereotyping can implicitly encourage or discourage individuals from doing specific tasks. Thus, the lack of gender parity in Wikipedia does not necessarily imply active sexism on the part of Wikipedians.

3. GENDERED WIKI-WORK

Wikipedia contributors do many tasks in order to keep the online encyclopedia operating at a high level. Wikipedians must not only write and edit content but also upload photos, create Wikipedia policies, and arbitrate disputes [8]. These are just a few examples of the many kinds of Wiki-work. However, even the primary category of writing and editing content hides many sub-categories. Editing a paragraph for grammar and reorganizing text to improve clarity, for example, are both forms of editing. Arguably, however, they constitute two distinct forms of practice with different skill and knowledge requirements, different relevant policies and quality criteria, and unique forms of reward and satisfaction. As a result, the different forms of writing and editing work on Wikipedia are unlikely to be interchangeable. Individual editors may gravitate towards certain types of work — for example those types which a Wikipedian feels best suit her skills or those which she feels are most important. There is already evidence suggesting that at least some Wikipedians specialize around specific social roles [11].

Examining gender distributions in different types of Wiki-work is essential for both theoretical and practical reasons. From a theoretical point of view, we know little about how factors such as gender may moderate engagement with online collaborative tools like Wikipedia. If we understand more about what types of work men and women are attracted to, we can turn a more discerning eye to the question of why that may be the case. What are the real or imagined characteristics of tasks that encourage certain individuals to gravitate towards them? A finer grained view of gender in Wiki-work could also help to reveal the influences of "occupational stereotypes." By observing these patterns and tracing their history we may learn more about how occupational stereotypes are formed. From a design point of view this exploration of gender is also essential. Wikipedia is invested in growing its user base and encouraging diversity. Answers to the question of who tends to do what kinds of Wiki-work could allow recruitment and educational efforts to be more focused. Understanding the role of gender may also help Wikipedia to do more than simply ask individuals to "contribute," and instead give potential contributors more specific ideas of the types of work they might like to do.

In this paper we focus on the activities of writing and editing articles. We then break these two high-level activities into more specific types of Wiki-work. As the foundation for our analysis we draw data about editing over time from Wikipedia's complete revision history. However, our questions about gender and Wiki-work require a more nuanced view of work than the Wikipedia API alone can provide. In addition to purely algorithmic analysis we also manually code revisions in order to determine the substantive type of editing work being done. As we have discussed above, there is ample reason to suspect that gender can influence not just how much work Wikipedians do or the topics on which they

work, but also the types of work (e.g. typo fixing, rephrasing for clarity, text reorganization) and how they do it.

We also compare the gender distribution of Wiki-work between Wikipedians who engage in different levels of activity. Like most online information systems, Wikipedia's pattern of participation follows a power-law function: a few people do most of the work while most people do very little. This indicates that the most active Wikipedians may have very different work practices than other Wikipedians. As a result it is important to examine potential gender differences separately for the most active Wikipedians.

4. METHOD

This research is part of a larger effort called the Wikipedia Progression of Participation (WPP) project, which aims to understand how new Wikipedians progress in their activities and attitudes over time. In this paper we focus on a sample of users who have recently created Wikipedia accounts. Examining the activities that individuals undertake as they begin their time as Wikipedians is crucial for understanding initial gendered work differences. This is a first step towards WPP's long-term goal of documenting the evolution of participation on Wikipedia over time.

The base population from which we draw our sample consists of 256,190 users who created a valid new account on the English-language Wikipedia between September 9th, 2010 and February 14th, 2011. Our analysis requires that we can determine the gender of each user. As a result, we limited our study to the 13,598 users (18.8%) who optionally declared a gender in their Wikipedia profile. Of these, 11,194 (82%) were men, and 2,402 (18%) were women. We have no way to verify users' gender, and it is likely that some accounts were used by more than one person of multiple genders. However, we have no reason to expect that such reporting errors systematically differ by gender.

From the sample of 11,194 gender-declaring Wikipedians we extracted a random stratified sample of 500 users. Five hundred is a statistically useful but practically manageable number of editors and revisions for our analysis. Because of the power law distribution of online participation, a simple random sample would contain mostly low-activity editors. Thus, we stratified our sample in order to equally capture all levels of engagement. We first observed the overall number of revisions each editor in the population made during her first 3 weeks with a Wikipedia account. From this distribution we extracted boundary thresholds for each quartile. The lower-boundary thresholds for each quartile were 0, 1, 2, and 4 total revisions respectively. We then randomly sampled an equal number of men and women from each quartile to build a total sample of 500 editors. For each sampled editor we extracted each revision made during an activity window of three weeks from the creation of the account. Three weeks of data were not available for 63 users in our sample, leaving a total of 437 users in our study.

4.1 Editing Typology

As a basis for our exploration of gender and Wiki-work we used the categorization of work proposed by Kriplean and colleagues [8]. Kriplean et. al defined 7 top-level categories which together contain 43 distinct types of work. We be-

Category	Revision Count					Number of Editors					Revision Size		
	Overall	Female		Male		Overall	Female		Male		Overall	Female	Male
Add Citations	9	5	(56%)	4	(44%)	8	5	(62%)	3	(40%)	972.4	190.8	1949.5
Add New Content	58	26	(45%)	32	(55%)	52	24	(46%)	28	(50%)	749.1	796.4	710.7
Change Wiki Markup	58	28	(48%)	30	(52%)	43	22	(51%)	21	(50%)	255.3	341.4	174.8
Create New Article	7	3	(43%)	4	(57%)	6	3	(50%)	3	(50%)	1589.4	2509.7	899.2
Delete Content	10	4	(40%)	6	(60%)	10	4	(40%)	6	(60%)	396.5	59.5	621.2
Fix Typo(s) / Grammar	16	9	(56%)	7	(44%)	15	8	(53%)	7	(50%)	120.8	16.7	254.6
Reorganize Existing Text	6	3	(50%)	3	(50%)	6	3	(50%)	3	(50%)	587.5	788.3	386.7
Rephrase Existing Text	39	20	(51%)	19	(49%)	33	16	(48%)	17	(50%)	250.2	387.2	105.9
Vandalism	10	6	(60%)	4	(40%)	8	5	(62%)	3	(40%)	708	987.8	288.2
Unsure	5	3	(60%)	2	(40%)	5	3	(60%)	2	(40%)	6.6	10	1.5
Overall	232	112	(48%)	120	(52%)	131	65	(50%)	66	(50%)	410.9	440.3	383.4

Table 1: Revision count, revision size, and number of editors by gender in the bottom 75% of our sample.

gan our typology with the 10 work types that were in the "Editing Work" category. We refined this typology through pilot testing and face validity agreement among the authors until we had a comprehensive list of meaningful types of editorial work that could be readily understood and identified by coders. Since we could not expect our coders to be expert editors, we provided comprehensive instructions and clarifying descriptions of the categories to make them as unambiguous and self-explanatory as possible. The final list included nine categories plus an "unsure" option. The main categories are listed in Table 1 and Table 2.

4.2 Coding & Data Aggregation

Our analysis requires that each revision be qualitatively coded and categorized according to the specific type of editing or writing work being done. We completed this task using Amazon's Mechanical Turk (MTurk). MTurk is an online labor market which was created to help complete large, piecemeal tasks. We paid MTurk workers 10 cents for coding blocks of 4 revisions, a task which could be completed in only 10-20 seconds.

Each revision in our sample was viewed and coded by at least three workers. Workers on MTurk were given thorough instructions as well as several "Gold" tasks in which we predefined correct answers and were able to communicate those answers as well as explanations to workers if they provided incorrect answers to those questions. The coding task followed a "check all that apply" model. Once all the data had been collected we applied a simple consensus model: if at least 2/3rds of the workers who coded each revision agreed on any given classification we accepted it. Finally, for each revision we also calculated the size of each revision using the standard metric Levenshtein distance. Levenshtein distance represents the minimum number of characters that one must add, delete, or change in order to create two identical strings.

5. RESULTS

Wikipedians in our sample made 4,549 revisions during the study period. Of these, 1,700 were revisions to Talk pages, User pages, and other non-article pages. In addition, 284 revisions are not reported because coders could not reach consensus on one or more substantive categories. Our total valid sample of coded revisions was 2,565. Table 1 and Table 2 illustrate the total number of revisions, the total number of unique Wikipedia editors, and the average size of each revision, divided by gender, for each type of editing work.

6. DISCUSSION & CONCLUSION

Our analysis provides a clearer and more detailed view of the "gender gap" by focusing not only on the number of male and female editors but also on the amount and type of Wiki-work that male and female Wikipedians tend to do. Similar to previous studies, we find that male Wikipedia editors drastically outnumber female editors overall. While the number of female editors (18%) was slightly larger than the number found in the 2010 UNU-MERIT study (13%), women remain under-represented overall.

However, the sheer number of editors of each gender does not tell the complete story. Several results are especially notable. First, among the bottom three quartiles of Wikipedians in our sample we found that men and women made similar numbers of revisions in nearly every category of Wiki-work. However, examining the top 25% of Wikipedians in our sample we observed a very different story. Here we found evidence that the overall number of revisions is far more skewed: just 27% of revisions were completed by women.

Examining the size of Wikipedia revisions reveals another interesting story. Among the bottom 75% of Wikipedians, women appeared to make larger revisions in many categories. However, it is difficult to observe a consistent pattern, especially because of the relatively small number of revisions in several categories. Only 9% of revisions were made by editors in the bottom 75% of our sample. As a result of the small sample and wide variation, no differences achieved statistical significance among the bottom 75%.

The pattern of revision size becomes stark, however, among the top 25% of Wikipedians. By examining the total number of revisions we observe that among the most active Wikipedians in our sample, women made far fewer revisions. However, we also observe that the most active women in our sample tended to make more sizeable revisions. The general pattern of results was highly significant overall (at $p < .001$) and consistent across most categories. Again, the relatively small number of revisions in several categories made it hard to achieve statistical significance. However, women made significantly larger revisions than men in the "Add New Content" and "Rephrase Existing Text" categories ($p \leq .05$).

We did not see evidence that men and women are attracted to different types of editing work. Of course, such differences may not exist. Alternatively, users may take time to grav-

Category	Revision Count						Number of Editors						Revision Size		
	Overall	Female		Male			Overall	Female		Male			Overall	Female	Male
Add Citations	111	30	(27%)	81	(73%)		47	14	(30%)	33	(70%)		314.5	542.9	230
Add New Content	477	162	(34%)	315	(66%)		93	43	(46%)	50	(50%)		380.8	583	276.8
Change Wiki Markup	793	231	(29%)	562	(71%)		107	48	(45%)	59	(60%)		153.3	190.1	138.1
Create New Article	39	10	(26%)	29	(74%)		24	9	(38%)	15	(60%)		1337.7	2019.8	1102.5
Delete Content	92	34	(37%)	58	(63%)		45	21	(47%)	24	(50%)		582.4	569.1	590.2
Fix Typo(s) / Grammar	126	40	(32%)	86	(68%)		54	23	(43%)	31	(60%)		275.2	489.4	175.5
Reorganize Existing Text	42	9	(21%)	33	(79%)		29	9	(31%)	20	(70%)		822.5	1046.2	761.5
Rephrase Existing Text	476	146	(31%)	330	(69%)		95	42	(44%)	53	(60%)		174.2	265.7	133.7
Vandalism	29	8	(28%)	21	(72%)		23	8	(35%)	15	(70%)		253	281.4	242.2
Unsure	61	20	(33%)	41	(67%)		42	15	(36%)	27	(60%)		77	23.8	102.9
Overall	2333	709	(30%)	1624	(70%)		124	57	(46%)	67	(50%)		246.6	349.4	201.7

Table 2: Revision count and size as well as total number of editors by gender in the top 25% of our sample.

itate towards specific types of work. If this were the case, we might not observe gender differences in an editor's first three weeks of participation. However, the analysis of revision size is another indication of gender differences in editing behavior. Notably, two areas of work in which women made significantly larger revisions involved creative production, synthesis, and reorganization of text. This finding adds to studies suggesting that, compared to men, women often develop more successful solutions to R&D problems posed through innovation brokers such as Innocentive [7].

It is important to note again that the nature of our chosen sample limits our ability to generalize. Since we sampled only new Wikipedians who began work relatively recently, we also cannot make claims about patterns over time. Finally, we have no information about whether those users who declare their gender in their profile are representative of Wikipedians as a whole. Despite these limitations, however, this analysis suggests that the story of Wikipedia's "Gender Gap" is perhaps not as straightforward as initial reports have suggested.

Our results can be viewed as both encouraging and discouraging for the effort to achieve greater gender parity on Wikipedia. On one hand, there was significant gender parity in number of revisions among the majority of Wikipedians — those in the bottom three quartiles of our sample. Furthermore, across quartiles women were better represented than the UNU-MERIT study would have led us to believe [4]. It is impossible to tell why this is the case. Wikipedia's efforts to attract more women may be bearing fruit, or there may be inaccuracies in prior studies. Several factors are likely at work. On the other hand, our results confirm that, compared to men, there are far fewer women editing Wikipedia. Furthermore, there was a particular gender skew in revision quantity among the Wikipedians who do most of the work. Problematically, the most active Wikipedians are also those who largely set policies, arbitrate disputes, and do other high-level tasks into which biases of worldview and temperament can subtly creep. This is the very group among which women may need more representation, but also one which may be more difficult to break into.

This study has just scratched the surface in exploring potential gender differences in Wiki-work. Although our results are largely descriptive and representative of a relatively small sample, we provide strong evidence that the "Gender Gap" story may be overly simplistic because it focuses entirely on the sheer number of male and female editors. Other recent recent has revealed the complex nature of Wikipedia's gender imbalances [9]. The more nuanced relationships between gender and Wiki-work we discuss are worthy of significant further study.

7. REFERENCES

[1] N. Cohen. Wikipedia ponders its Gender-Skewed contributions. *The New York Times*, Jan. 2011.

[2] R. M. Dawes, A. J. V. D. Kragt, and J. M. Orbell. Not me or thee but we: The importance of group identity in eliciting cooperation in dilemma situations: Experimental manipulations. *Acta Psychologica*, 68(1-3):83–97, Sept. 1988.

[3] S. Gardner. Nine reasons women don't edit wikipedia. http://suegardner.org/nine-reasons-why-women-dont-edit-wikipedia-in-their-own-words/, 2011.

[4] R. Glott, R. Ghosh, and P. Schmidt. Analysis of wikipedia survey. topic: Age and gender differences. UNU-MERIT, 2010.

[5] L. S. Gottfredson. Circumscription and compromise: A developmental theory of occupational aspirations. *Journal of Counseling Psychology*, 28(6):545–579, 1981.

[6] D. A. Huffaker and S. L. Calvert. Gender, identity, and language use in teenage blogs. *Journal of Computer-Mediated Communication*, 10(2), 2006.

[7] L. B. Jeppesen and K. R. Lakhani. Marginality and problem-solving effectiveness in broadcast search. *Organization Science*, 21(5):1016–1033, 2010.

[8] T. Kriplean, I. Beschastnikh, and D. W. McDonald. Articulations of wikiwork: uncovering valued work in wikipedia through barnstars. In *Proc. CSCW*, pages 47–56, San Diego, CA, USA, 2008. ACM.

[9] S. Lam, A. Uduwage, Z. Dong, S. Sen, D. Musicant, L. Terveen, and J. Riedl. WP:Clubhouse? an exploration of wikipedia's gender imbalance. In *Proc. Wikisym 2011*, Mountain View, CA, 2011. ACM.

[10] L. McDowell. *Capital Culture: Gender at Work in the City*. Wiley-Blackwell, Dec. 1997.

[11] H. T. Welser, D. Cosley, G. Kossinets, A. Lin, F. Dokshin, G. Gay, and M. Smith. Finding social roles in wikipedia. *Proceedings of the 2011 iConference*, page 122–129, 2011.