Reducing Bias in Wikipedia's Coverage of Political Scientists

INTRODUCTION

One in five political scientists with biographies on Wikipedia are women, while nearly half are

American. Biases on Wikipedia can cause real harm, so I created or expanded a political science-related

Wikipedia article every day for a year, focusing on writing new pages about political scientists from

underrepresented groups. In this piece I show that Wikipedia's coverage of political scientists remains

skewed by gender and nationality, and I suggest ways for political scientists to improve Wikipedia's

representation of the discipline.

Anyone can write an article on Wikipedia, including a biography of another person, but an

overwhelming proportion of these biographies are about men from wealthy countries. This has

prompted a kind of digital activism: for nearly as long as Wikipedia has existed, there have been

collective efforts to make its coverage of people fairer (Redden 2016).[i] Almost a decade of work has

focused specifically on increasing the proportion of biographies about women in science (Wade and

Zaringhalam 2018). But while thousands of new pages have been written about underrepresented

natural scientists, medical researchers, and engineers, Wikipedia's treatment of social scientists has not

received the same level of attention.

And yet, Wikipedia's coverage of social sciences suffers from the same problems as its coverage

of natural and medical sciences. Adams, Brückner, and Naslund (2019) have found that sociologists who

are men or who are white are disproportionately likely to be the subject of a biography on Wikipedia,

while Schellekens, Holstege, and Yasseri (2019) identified similar gender gaps among pages about economists, physicists, and philosophers. As Alter *et al.* (2020) have argued, and Ackerly and Michelitch (2021) discuss in this symposium, there is reason to expect that these biases would also apply to political scientists. In order to match the estimated proportion of women among full-time political scientists in the United States, Wikipedia's biographies of women in political science would need to appear at nearly double the overall proportion of Wikipedia biographies about women (Fraga *et al.* 2011). [ii]

These sorts of biases echo enduring problems in political science itself. Gender gaps have been consistently identified in political science citations, syllabi, and books, while the academic study of politics has historically focused on a small number of wealthy countries (Dion and Mitchell 2019; Wilson and Knutsen 2020). The prominent Women Also Know Stuff and People of Color Also Know Stuff initiatives were partly motivated by a closely related bias: publications and events about politics being less likely to include experts who are white women or people of color (Beaulieu *et al*. 2017; Lemi *et al*. 2019).

Like other types of bias in the discipline, biased coverage on Wikipedia can have serious effects. In a typical month, Wikipedia articles about politics are viewed hundreds of millions of times. [iii] Experiments have demonstrated that text on highly viewed Wikipedia pages can affect outcomes from the way that scientific topics are described in academic papers to the revenue of tourist destinations (Hinnosaar *et al.* 2017; Thompson and Hanley 2017). Researchers have specifically argued that biased coverage of a profession on Wikipedia can discourage people from joining that profession, because of an absence of visible role models (Hinnosaar 2019; Wade and Zaringhalam 2018).

To address these issues, I undertook a year-long project to reduce bias among political science-related Wikipedia articles, focusing on gender and regional disparities in biographies about political scientists. Every day of 2020, I created or expanded a Wikipedia article about a political scientist from an

underrepresented group. So how productive were these efforts, and how much work remains before Wikipedia biographies of political scientists are a fair representation of the field?

First, I describe the distributions of gender and nationality among Wikipedia pages about political scientists. Then, I discuss the limitations of an individual effort like mine that aims to address a systemic problem. Finally, I provide concrete suggestions for how others can contribute.

BIASED COVERAGE

To measure the extent of bias in pages about political scientists, it is necessary to classify Wikipedia's biographies of political scientists according to salient characteristics.

Following past literature on Wikipedia's bias, I will focus on two: gender and nationality. These are certainly not the only interesting attributes, but they have the advantage of being straightforwardly tracked across Wikipedia pages; other characteristics, like ethnicity, can be even more contingent on time and place. By "bias", I specifically mean that the distribution of Wikipedia articles about political scientists by gender and nationality does not resemble the real distributions of political scientists by gender and nationality. Brown (2011) has found that Wikipedia's coverage of politics "suffers less from inaccuracies than omissions", and in this vein I study the omissions in its coverage of political scientists, rather than focusing on issues in the contents of its pages. I make no other inferential or associational claims about these biases.

There are only a few thousand pages about political scientists on Wikipedia, so exact manual counts are possible. To identify the distribution of political scientists by gender, I manually checked every biography that was classified as being about a political scientist, and identified whether or not the article subject is a woman (classifying article subjects by gender is not trivial; as I discuss in an online

appendix, the precise feature I classified is whether or not a page subject is referred to by she/her

pronouns).[iv] In late 2020, there were 3,335 biographies of political scientists on the English Wikipedia,

but only 670 of these (about 20%) were biographies of women. This is substantially lower than estimates

of the proportion of women in the discipline, which (for example) in the United States is close to 30%

(Fraga *et al.* 2011; Nelson 2017).

What about nationality? The left part of Figure 1 maps the distribution of political scientists with

English Wikipedia biographies by country. Naturally, we might expect political scientists from English-

speaking regions to be better-represented on an English-language encyclopedia. So I counted the

nationality of every political scientist across all 49 languages in which Wikipedia has at least one

biography of a political scientist. I identified 11,077 biographies of political scientists in any language,

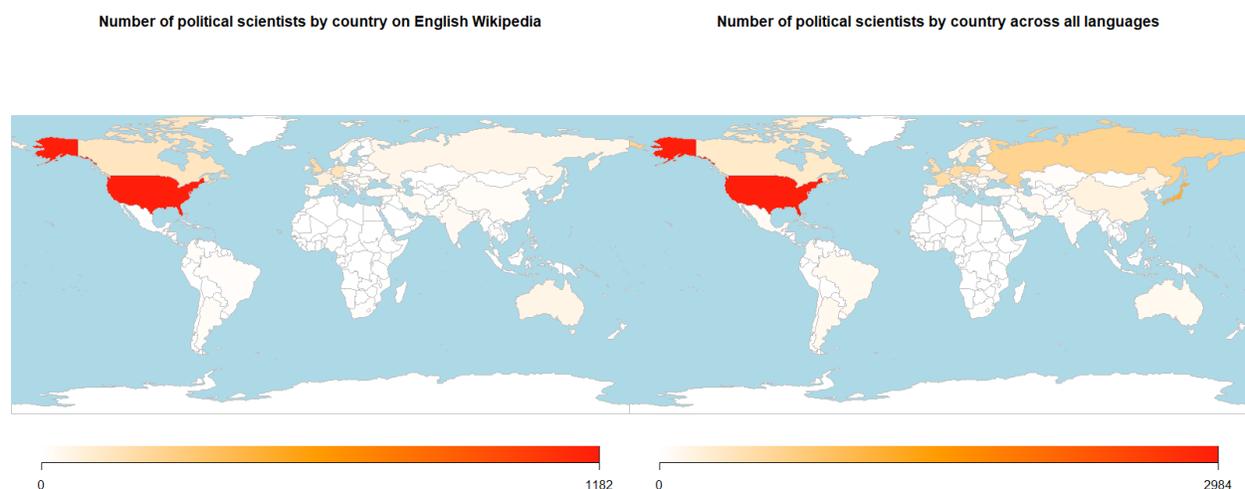and their nationalities are mapped on the right of Figure 1.

Figure 1: The distribution by nationality of political scientists who have a biography on the Wikipedia in English (left), and in any

language (right).

In the distribution of English-language pages, the United States is an overwhelming outlier:

about 43% of all the biographies of political scientists on Wikipedia are about Americans. Canada, the

United Kingdom, Germany, and France are also highly represented, and there is sparse coverage of political scientists outside of North America or central and western Europe. The major difference when including non-English pages is that Japan becomes dramatically better-represented, as do Russia and some Eastern European countries.

There are two reasons to believe that this distribution does not resemble the real distribution of political scientists around the world.

First, it does not agree with other estimates. Ideally we would compare the distribution of Wikipedia articles to some gold standard estimate of the distribution of political scientists by nationality, but this distribution has eluded confident identification. However, estimates have been made of the number of political scientists within a specific country or region at various times. Klingemann (2008) estimated that there were about 10,000 political scientists in Europe, while around the same time the APSA faculty database identified over 9,000 political scientists in the United States (Fraga *et al.* 2011). The Wikipedia data do not quite reflect these numbers, since there are hundreds more American political scientists than European political scientists on English Wikipedia. And if these figures are correct, then Wikipedia's proportion of American political scientists is only representative if the entire rest of the world contains about 5,000 political scientists. Analyses of political science activity in other places, from political science instruction in "several thousand colleges" in India (Shah 2001) to the volume of articles written by Brazilian political scientists (Nicolau and Oliveira 2017), suggest that this is not a reasonable estimate. Evidently, Wikipedia's coverage of political scientists is biased towards the United States.

A second problem is that several countries with active political science communities have almost no representation. Figure 2 classifies countries by whether or not at least one political scientist of that nationality has a biography on Wikipedia. Of the 85 UN member states with no political scientists on the

English Wikipedia, at least 9 have a national political science association or journal. And several other

countries with large political science communities, like Brazil, India, Japan, and Nigeria, have only a few

dozen political scientists on Wikipedia.

Regions where at least one political scientist has a page

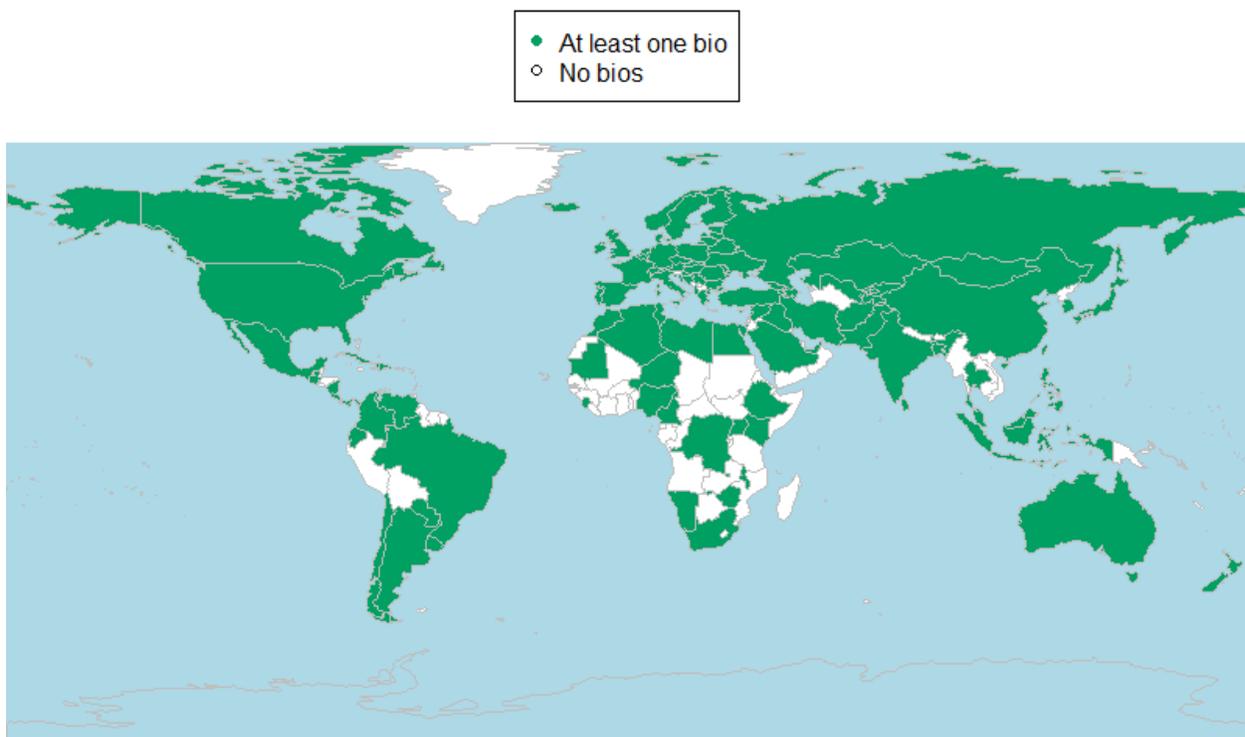- At least one bio
- No bios

Figure 2: Countries that have at least one political scientist with a biography on English Wikipedia.

A final hint of missing political scientists is that many countries have large numbers of Wikipedia

pages about political scientists in other languages, but very few in English. The most extreme case is

Japan: there are more than 1,000 Japanese political scientists with Japanese Wikipedia biographies, but

only a few dozen Japanese political scientists have English Wikipedia biographies.

An exact count demonstrates that Wikipedia is disproportionately likely to include political scientists who are men, and political scientists who are Americans. This disparity is a form of bias, and is consistent with the biases identified by similar analyses.

PROPOSED SOLUTIONS

The fact that only about one in every five biographies of political scientists on the English Wikipedia are about women, while nearly half are about Americans, holds true after I made daily efforts to reduce these biases for a year. These efforts involved writing 297 new pages, which were viewed about 300,000 times during 2020.[v] This included about 260 pages about women, and added about a dozen countries to the list of nationalities with at least one political scientist on Wikipedia. I also expanded about 70 existing pages. But the population of Wikipedia articles is so large that even sustained individual efforts make little difference. Narrowly construed, about 250 of the articles I wrote were new pages about political scientists (others were articles on closely related topics, such as a new article about the subfield of gender and politics). So approximately just 7% of Wikipedia's articles about political scientists were created during this project. My contributions also hardly addressed, and in some ways may even have exacerbated, the bias by nationality. Evidently, collective action is needed.

However, there are several obstacles that prevent people who have never written a Wikipedia article from helping to rapidly fix biases. Though there are many guides to the mechanics of writing Wikipedia articles, the complexities of writing a Wikipedia biography, particularly of a living academic, are more arcane. On the one hand, Wikipedia has complicated rules about what can be included in biographies of living people, largely motivated by concerns about copyright and libel. It also has rules about which academics are considered "notable" enough to have an article; academics who are judged to not satisfy these rules might have their page deleted (and these deletions may exacerbate existing

biases: see Tripodi 2021).[vi] On the other hand, the interests of the page subject need to be considered;

their page must be accurate, useful, and protected from vandalism (though, if needed, pages which

were mostly written by one author can be quickly deleted at the request of that author). I therefore

suggest three approaches for new editors which are likely to produce pages that are safe, useful,

counter to the website's bias, and motivated by past research (Luo, Adams, and Brückner 2018;

Schellekens, Holstege, and Yasseri 2019).

First, many of these considerations are nullified if the page subject is deceased. Information

about many important political scientists of previous generations can only be found by searching

through paywalled journals and reading paywalled obituaries. By writing pages about historical political

scientists on Wikipedia, we make information about their work freely available. Many underrepresented

political scientists from previous generations are still missing from Wikipedia, so this can be

accomplished in large-scale initiatives. One model could follow the example of Rachael Reavis's lab, in

which students write Wikipedia pages about underrepresented psychologists, and could incorporate

ideas from other works in this symposium about using Wikipedia editing as a project for students

(Michelitch and Wilfahrt 2021; Norell 2021; Sengupta and Ackerly 2021; Weingartner 2020).

A second strategy is to translate Wikipedia pages from other languages into English. This can be

easier than writing a page from scratch, and partially addresses concerns about writing biographies of

living people: it is more likely that the page subject will be judged notable, and translators can select

pages that appear particularly high-quality.

Finally, though I have focused in this piece on the bias among biographies of political scientists,

a related problem is bias in citations. Introducing citations is comparatively fast and simple, and

egregious disparities in citations are easy to find. One need look no further than Wikipedia's page on

"political science", which is the top Google result for "political science", and receives about 700,000

views a year.[vii] In late 2020, the first authors in its reference and suggested reading sections included 26 men and 1 woman.[viii]

CONCLUSION

At a rate of one contribution per day, one person might take about 2 more years to increase the proportion of Wikipedia biographies about women in political science up to its current proportion in the field. Even more work would be needed to improve coverage of political scientists from underrepresented countries, or to address biases in citations. Worse, if we do not change the imbalances among new pages created, then the proportions of pages will tend back to their original skew; we can only make a long-term difference if we permanently reduce the bias among new pages. However, any efforts to address these problems need to be directed towards effective and useful additions. New pages need to not be deleted from Wikipedia, and must provide a positive service to the encyclopedia, to its readers, and especially to the page's subject. Three promising next steps are for any interested person to write pages about political scientists from the history of the discipline, translate pages that already exist in other languages, and introduce citations to political scientists who are, at present, severely underrepresented in the world's most important reference work.

WORKS CITED

Ackerly, Brooke and Kristin Michelitch. 2021. "Wikipedia and Political Science: Addressing Systematic Biases with Student Initiatives". *PS: Political Science and Politics*.

Adams, Julia, Hannah Brückner, and Cambria Naslund. 2019. "Who Counts as a Notable Sociologist on Wikipedia? Gender, Race, and the 'Professor Test'". *Socius: Sociological Research for a Dynamic World.* 5: 1–14. doi: 10.1177/2378023118823946

Alter, Karen J., Jean Clipperton, Emily Schraudenbach, and Laura Rozier. 2020. "Gender and status in American political science: Who determines whether a scholar is noteworthy?" *Perspectives on Politics*. 18(4): 1048–1067 doi: 10.1017/S1537592719004985

Beaulieu, Emily, Amber E. Boydstun, Nadia E. Brown, Kim Yi Dionne, Andra Gillespie, Samara Klar, Yanna Krupnikov, *et al*. 2017. "Women Also Know Stuff: Meta-level mentoring to battle gender bias in political science". *PS: Political Science and Politics* 50(3): 779–783. doi: 10.1017/S1049096517000580

Brown, Adam R. 2011. "Wikipedia as a data source for political scientists: Accuracy and completeness of coverage." *PS: Political Science and Politics* 44(2): 339-343.

Dion, Michelle L. and Sara McLaughlin Mitchell. 2019. "How Many Citations to Women Is 'Enough'? Estimates of Gender Representation in Political Science." *PS: Political Science and Politics* 53(1): 107–113. doi: 10.1017/S1049096519001173

Fraga, Luis Ricardo, Terri E. Givens, Dianne M. Pinderhughes, Manuel Avalos, David Covin, Frances Hagopian, Lisa García-Bedolla, *et al*. 2011. "Political Science in the 21st Century: Report of the Task Force on Political Science in the 21st Century." American Political Science Association. https://www.apsanet.org/portals/54/Files/Task%20Force%20Reports/TF_21st%20Century_AllPgs_webres90.pdf

Hinnosaar, Marit, Toomas Hinnosaar, Michael E. Kummer, and Olga Slivko. 2017. "Wikipedia Matters". SSRN Scholarly Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3046400

Hinnosaar, Marit. 2019. "Gender inequality in new media: Evidence from Wikipedia". *Journal of Economic Behavior & Organization* 163: 262–276. doi: 10.1016/j.jebo.2019.04.020.

Klingemann, Hans-Dieter. 2008. "Capacities: Political science in Europe." *West European Politics* 31(1–2): 370–396. doi: 10.1080/01402380701835181

Lemi, Danielle Casarez, Maricruz Osorio, and Tye Rush. 2019. "Introducing People of Color Also Know Stuff." *PS: Political Science and Politics* 53(1): 140–141. doi: 10.1017/S1049096519001069

Luo, Wei, Julia Adams, and Hannah Brückner. 2018. "The Ladies Vanish? American Sociology and the Genealogy of its Missing Women on Wikipedia." *Comparative Sociology* (17): 519–56. doi: 10.1163/15691330-12341471.

Michelitch, Kristin and Martha Wilfahrt. 2021. "Improving Open-Source Information on African Politics, One Student at a Time". *PS: Political Science and Politics*.

Norell, Elizabeth. 2021. "Civic Engagement Meets Service Learning: Improving Wikipedia's Coverage of State Government Officials". *PS: Political Science and Politics*.

Nicolau, Jairo and Lilian Oliveira. 2017. "Political science in Brazil: An analysis of academic articles (1966–2015)". *Sociologia & Antropologia* 7(2). doi: 10.1590/2238-38752017v723

Nelson, Donna J. 2017. "Diversity of Science and Engineering Faculty at Research Universities." In *Diversity in the Scientific Community Volume 1: Quantifying Diversity and Formulating Success*, eds. Donna J. Nelson and H. N. Cheng. Washington, DC: American Chemical Society.

Redden, Molly. 2016. "Women in science on Wikipedia: will we ever fill the information gap?" *The Guardian*. https://www.theguardian.com/technology/2016/mar/19/women-in-science-on-wikipedia-will-we-ever-fill-the-information-gap

Schellekens, Menno H., Floris Holstege, and Taha Yasseri. 2019. "Female scholars need to achieve more for equal public recognition". ArXiV. https://arxiv.org/abs/1904.06310

Sengupta, Anasuya and Brooke Ackerly. 2021. "Wikipedia edit-a-thons: sites of struggle, resistance, and responsibility". *PS: Political Science and Politics*.

Shah, Ghanshyam. 2001. "Political science in India: A discipline and intellectual pursuit." *The Indian Journal of Political Science* 62(1): 11–23.

Thompson, Neil C. and Douglas Hanley. 2018. "Science Is Shaped by Wikipedia: Evidence From a Randomized Control Trial". MIT Sloan Research Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3039505

Tripodi, Francesca. 2021. "Ms. Categorized: Gender, notability, and inequality on Wikipedia". *New Media & Society* 1–20. doi: 10.1177/14614448211023772

Wade, Jess and Maryam Zaringhalam. 2018. "Why we're editing women scientists onto Wikipedia". *Nature Careers*. https://www.nature.com/articles/d41586-018-05947-8

Weingartner, Tana. 2020. Wikipedia Lacks Profiles Of Women. These College Students Are Changing That. *Cincinnati Public Radio*. https://www.wvxu.org/post/wikipedia-lacks-profiles-women-these-college-students-are-changing#stream/0

Wilson, Matthew Charles and Carl Henrik Knutsen. 2020. "Geographical coverage in political science research." *Perspectives on Politics* 1–16. doi: 10.1017/S1537592720002509

[i] https://en.wikipedia.org/w/index.php?title=Wikipedia:WikiProject_Countering_systemic_bias&oldid=6332105

[ii] https://en.wikipedia.org/w/index.php?title=Wikipedia:WikiProject_Women_in_Red&oldid=995690733

[iii] https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Politics/Popular_pages

[iv] https://samuelbaltz.net/supplements.html

[v] https://pageviews.toolforge.org/userviews/?project=en.wikipedia.org&platform=all-access&agent=user&namespace=0&redirects=0&range=last-year&sort=views&direction=1&view=list&user=Astrophobe

[vi] https://en.wikipedia.org/w/index.php?title=Wikipedia:Notability_(academics)&oldid=990432229

[vii] https://pageviews.toolforge.org/?project=en.wikipedia.org&platform=all-access&agent=user&redirects=0&range=all-time&pages=Political_science

[viii] https://en.wikipedia.org/w/index.php?title=Political_science&oldid=977224980