The Gender Balance Assessment Tool (GBAT): A Web-Based Tool for Estimating Gender Balance in Syllabi and Bibliographies

Jane Lawrence Sumner, *University of Minnesota, Twin Cities*

ABSTRACT   This article introduces a web-based tool that scholars can use to assess the gender balance of their syllabi and bibliographies. The citation gap in political science is described briefly as well as why under-citing women relative to men is a problem that should be addressed by the field. The Gender Balance Assessment Tool (GBAT) is presented as a way to make assessing gender balance easier with the aim of remedying the gender gap. This is followed by an outline that explains in nontechnical terms how the tool identifies author names and then predicts their gender to produce a single document-level percentage of women authors. Finally, best practices for diversity in syllabi and bibliographies are discussed, and various public sources that can be used to find scholarly work by women, as well as scholars of color, are listed.

Women also know stuff, but we might not know that from reviewing political science bibliographies. Work by women is far less likely to be cited than similar work by men (Maliniak, Powers, and Walter 2013), and men are particularly unlikely to cite women (Mitchell, Lange, and Brus 2013).

Women are under-cited as well as under-assigned. For example, in the top 200 most-frequently assigned works in the Open Syllabus Project's "Politics" section,[1] only 15 works are authored by at least one woman, whereas 20 are authored by at least one man named Robert. Of the 219 total authors on that list, 204 are men and only 15 (6.8%) are women, far less than in the discipline as a whole. More rigorous analyses have found that women authors may appear as often in international relations syllabi as they do in the field. However, this may be driven by women being more likely to assign other women, which is indicative of under-assignment by men (Colgan 2015).[2]

Because researchers tend to cite, at least initially, the works assigned to them in coursework (Nexon 2013), the gender gap in citations is exacerbated by under-assignment in syllabi. Repercussions of the citation gap exceed a normative desire for descriptive diversity: decisions about hiring, promotion, tenure, and raises often are informed by citation counts.

Many other explanations are offered for why women are not cited as frequently as men. For example, women are less likely to cite themselves in their own research (Colgan 2015; Maliniak, Powers, and Walter 2013; Mitchell, Lange, and Brus 2013). Another explanation is that scholars tend to be most familiar with work by people within their social networks, which tend to be gendered (Mansbridge 2013).

A third explanation is that assessing gender balance in bibliographies and syllabi can be difficult and tedious.[3] Determining the percentage of the 200 most-assigned works from the Open Syllabus Project that were by women authors involved researching many unfamiliar names, determining which identified as women, and summing the total number of authors. The process took 20 to 30 minutes. Although this may not seem excessive, those who otherwise might be inclined to assess their gender balance may view this process as an impediment.

This article introduces a web-based tool that I created to help scholars assess the gender balance of their bibliographies and syllabi.[4] Whereas many scholars have long assessed the gender balance manually, far more have not. This tool makes this process fast and easy for people not already predisposed to manual assessment. It uses RShiny to implement an algorithm that identifies author names, probabilistically codes each author's gender, and then provides the user with an estimate of the percentage of authors who are women. This process is less accurate than hand-coding, but it is much faster and easier and provides users with a fairly reliable and accurate estimate. For instance,

---

Jane Lawrence Sumner *is assistant professor in political science at the University of Minnesota, Twin Cities. She can be reached at jlsumner@umn.edu.*

 doi:10.1017/S1049096517002074

when applied to the 200 most-frequently assigned politics texts, the tool identified 211 names and determined that 9.68% were women. This compared to 218 names and 6.8% when hand-coding. Although the tool found fewer authors and a larger percentage of women than was produced by hand-coding, the result was similar, much faster, and a huge improvement over not assessing gender balance at all. The sources of this inaccuracy are discussed in more detail herein.

The next section explains how the tool works and describes in detail how it identifies names and probabilistically codes gender. The section that follows briefly discusses the two most frequently asked questions: (1) What proportion of women should be the goal?; and (2) How should scholars aim to balance their bibliographies and syllabi?

THE GBAT

The GBAT makes estimating diversity quick and easy, to help those who would like to but who do not otherwise undertake such an assessment. It works by identifying author names in a document and then estimating the probability that an author is a woman. The tool then aggregates each probability to approximate the percentage of women authors in the list. The entire process, from uploading to final estimate, typically takes less than a minute.

The following sections explain the two primary components of the algorithm: identifying likely author names and computing gender probabilities.

look like names." This algorithm works by removing character strings that are unlikely to be names and then by identifying strings of characters that follow patterns that "look like" names.

Before identifying names, the algorithm removes words, letters, and symbols that are unlikely to appear in names: most conjunctions and stop words (leaving in "I," "a," and "and"); numbers; and words and word stems from a list of common title, journal, and publisher names (table 1). The algorithm replaces these characters with spaces. The following example of a resulting text has fewer words and more empty spaces:

> ". Power Shift. Matthews, Jessica T. Affairs. . active ying. Austen-Smith, David, Wright, John R. an cal . . True Clash izations. Inglehart, Ronald, Norris, Pippa. y."

Stray punctuation marks and repeated spaces are deliberately retained because they make it easier to identify names. This is important because some of the characteristics that allow us to identify that a string of characters "looks like" a name—two or three words in a row all beginning with capital letters, for instance—also are shared by other words in titles and journal names. Removing words, inserting spaces, and retaining punctuation prevents many titles and journal names from falsely being identified as names. From the resulting text, regular expressions and the R package openNLP[7] can be used to extract a list of probable names.

Because researchers tend to cite, at least initially, the works assigned to them in coursework (Nexon 2013), the gender gap in citations is exacerbated by under-assignment in syllabi.

IDENTIFYING NAMES

Whereas computers excel at implementing written directions faster and more consistently than humans, many tasks that can be done easily by humans are difficult for computers.[5] Identifying names as distinct from other non-name words is one such task. For example, consider the last three entries on the top 200 list, which resemble entries in a bibliography or a syllabus:

> Power Shift. Matthews, Jessica T. Foreign Affairs.
> Counteractive Lobbying. Austen-Smith, David, Wright, John R. American Journal of Political Science.
> The True Clash of Civilizations. Inglehart, Ronald, Norris, Pippa. Foreign Policy.

Most humans can quickly identify that this list contains five authors, two of which are women. Computers do not recognize the nuanced human idea that some words "just look like names" and therefore this intuition does not translate easily into code. Computers can be programmed to identify patterns that resemble names but are seldom as precise as humans.[6] For that reason, enlisting a computer would not be sensible if a syllabus or a bibliography were actually as short as the previous example. However, if the list of citations were much longer, the computer's speed and untiring repetition would become useful, even with the slight cost of the accuracy of human coding. The key is telling a computer what it means that some words "just

In the previous example, the tool identifies "Power Shift," "Mathews, Jessica," "Austen-Smith, David," "Wright, John," "True Clash," "Inglehart, Ronald," and "Norris, Pippa" as probable author names. This is good for a first pass; it correctly identifies all author names and contains only two false names. With this list of probable author names, the tool moves on to the next stage—probabilistically predicting gender—which also will eliminate many of the false-positive names.

PROBABILISTIC PREDICTION OF GENDER

Most academics likely can quickly identify in this list of probable names that three first names are common among men (i.e., David, John, and Ronald); two first names are typically associated with women (i.e., Jessica and Pippa); and two phrases are not names (i.e., Power Shift and True Clash).[8] Computers lack the inherent human ability to make these same contextual judgments; fortunately, algorithms have been written to help them do so.

The GBAT predicts an author's gender from the author's given name using the genderize.io algorithm, as implemented in the genderizeR package for R (Wais 2015). Unlike other data sources—such as US Social Security Administration data, which only includes data on names that are common in the United States—genderize.io and genderizeR use social-media data. Therefore, it can predict gender for many more names, allowing for greater inclusion. An additional benefit of this tool is that it often screens out the non-names included in the probable-names list. If the algorithm determines that the first

word of the non-name term (e.g., "Power" or "True") is probably not a name because there is insufficient data to predict the gender, it is omitted from the gender prediction.[9] The tool then aggregates the name-specific probabilities, producing an overall percentage of authors that are likely to be women.[10]

However, there are shortcomings to this approach. First, authors identified only by initials cannot be categorized using this method and are dropped from the estimation process. Second, because the tool aggregates probabilities rather than

Mancur Olsen's gender cannot be predicted due to insufficient data; works by Dani Rodrik (probability = 0.61) and Alexis de Toqueville (probability = 0.48) incorrectly inflate the proportion of probable women; and work by Lee Epstein (probability = 0.25) incorrectly decreases it. The result is fewer total authors identified and an inflated estimate of the percentage of women. As a result, the tool usually provides an estimate that closely resembles reality, but users must be aware that it is only an estimate; particular characteristics of their documents may lead to more

This tool was developed with modest intentions. Rather than produce the most accurate estimate of diversity within a syllabus, the aim is to make assessing gender diversity so easy and quick that more scholars will do it.

dichotomous designations of "man" and "woman," names that are common among both genders can throw off the estimate. Third, the algorithm cannot predict gender for names that are particularly uncommon worldwide due to lack of data; these names also are dropped. Bibliographies and syllabi with high levels of any of these three issues will be less accurate. For instance, in the Open Syllabus Project example referred to previously, W. W. Rostow, V. O. Key, and four other names with initials only are dropped;

or less accuracy. This highlights a key tradeoff of the tool: it will never be as accurate as thorough hand-coding and should not be used as a replacement for it. However, it is a quick and easy estimation tool for those not already predisposed to hand-coding.

DISCUSSION

This tool was developed with modest intentions. Rather than produce the most accurate estimate of diversity within a syllabus, the aim is to make assessing gender diversity so easy and quick that more scholars will do it. It is hoped that this will lead to (1) a gradual decrease in the gender gap as scholars realize the degree of their under-citation of women; and (2) a rethinking of what is being cited, why, and what is being overlooked. A low percentage of women should be an invitation to explore what other material exists and may be unintentionally excluded.[11] However, this raises two important questions, the first of which pertains to best practices and the second to how the descriptive diversity of our bibliographies and syllabi can be improved.

The modal question asked in response to this tool is: "My bibliography/syllabus was N% women; is that good?" Reasonable minds differ on this point; however, whether a particular percentage is normatively desirable depends—at a minimum—on the topic at hand. Some subfields (e.g., political methodology) have far less diversity than others, including Race, Ethnicity,

Table 1
Common Title Words

Africa	Develop	Justice	Quarter	Vote
Americ	E.U.	Knowledg	Reconsider	Voting
Analy	Eastern	Legislat	Relev	Washington, DC
Asia	Econo	Little, Brown	Research	What
Attitud	Elect	Local	Respon	When
Basic	Essay	Logic	Review	Where
Bayes	Estimat	Majorit	School	Who
Behavi	EU	Maximi	Science	Why
Behavio	Europe	Maximu	Second	Yale
Book	Explor	Method	Social	
Bureau	First	Michigan	Sociali	
California	Forthcoming	Minorit	Societ	
Cambridge	Framework	Model	State	
Capital	Geograph	Multipl	Statistic	
Central	Global	Nation	Studies	
Civil	Govern	New York	Study	
Collect	Harvard	of Georgia	Theor	
College	Hypoth	of Virginia	Transit	
Communis	Improv	Oxford	Typescript	
Communis	Institut	Paper	U.K.	
Congress	Interact	Polic	U.S.	
Control	International	Politi	UK	
Critic	Introduc	Press	Understand	
Democ	Journal	Problem	Univers	
Depart	Judici	Publi	US	

and Politics. Imposing uniform standards across subfields may not be sensible: a quantitative political methodology syllabus that has 20% women authors may be representative of the diversity in the subfield, whereas the same 20% on a syllabus about women in politics would be extremely problematic.[12] If scholars are unsure about the diversity of their subfield, reports such as the American Political Science Association (APSA) Status of Women in the Profession provide details about diversity in the field as a whole. Other professional organizations, such as the International Studies Association (ISA), publish similar information. Organized sections within APSA, the Midwest Political Science Association (MPSA), and ISA also may want to publish descriptive data about their membership to determine how relatively diverse their bibliographies and syllabi are.

The second most frequently asked question concerns how to make bibliographies and syllabi more diverse: How and where do we find relevant articles by women? Fortunately, the answer in many cases is simple: the Internet is replete with information about women and their research interests. Notably, the website WomenAlsoKnowStuff[13] exists to address this problem and maintains a list of women scholars who create profiles on its website (Beaulieu et al. 2016). The website is organized by subject area, providing a list of women scholars in many subfields. Subfield-specific women's groups also maintain public membership lists, including Visions in Methodology[14] (political methodology), Women in Conflict Studies[15] (conflict), and Journeys in World Politics[16] (international relations).

Although this article focuses on gender, scholars of color likely face many of the same problems as women regarding citation and assignment gaps, and scholars should be equally mindful of racial diversity when assessing bibliographies and syllabi. There are many resources on the Internet for locating scholars of color. The Twitter account @PoCAlsoKnow amplifies accomplishments by people of color in academia (@PoCAlsoKnow 2016).[17] The APSA Latino Caucus maintains a public membership list that is organized by subfield.[18] Similarly, the APSA Asian Pacific American Caucus maintains a public list that, although not organized by subfield, includes scholars' research interests.[19] The National Conference of Black Political Scientists does not maintain a public membership list but allows the public to search for members. Using the Advanced Search feature,[20] users can select from a list of member types and search by field and subfield on the next page.

CONCLUSION

Although women actively contribute research in political science, their work is cited less frequently than their male counterparts. Their work also tends to be underrepresented on syllabi, which may exacerbate the citation gap. This gender gap has deleterious effects for women scholars because citation counts affect decisions about hiring, tenure, promotion, and raises.

Yet, even though the disadvantages of the citation gap are well known, this knowledge is only part of the battle. Scholars must keep these issues in mind when citing research or constructing syllabi. Although many scholars hand-code their bibliographies and syllabi, many more do not. For those not already predisposed to assessing descriptive diversity in their citations, determining the gender balance of a particular bibliography or syllabus may be viewed as tedious, time-consuming, and difficult. To make the task faster and easier, this article describes a web-based tool that allows users to easily upload a bibliography or syllabus and,

within a minute, receive a probabilistic estimate of their bibliography or syllabus gender balance.

My goal in presenting this tool is to remove many of the practical roadblocks that deter scholars from assessing the gender balance of their syllabi and bibliographies. My hope is that this will encourage more people to assess the gender balance and then use this information to understand why they are citing particular authors and articles and whether they can make their syllabi and bibliographies more representative of the diversity of the field as a whole. To that end, this article also highlights resources to easily find women scholars and scholars of color who are conducting research in their field.

NOTES

1. The Open Syllabus Project is a collection of more than a million undergraduate and graduate syllabi from approximately the last 10 years, collected by crawling and scraping public university websites. The "Politics" section contains 99,133 syllabi from colleges and universities in the United States, the United Kingdom, Canada, Australia, and New Zealand. Available at http://explorer.opensyllabusproject.org (accessed September 1, 2016).

2. Not assigning research by women also may have negative consequences for students (Carrell, Page, and West 2010; Cassese, Bos, and Duncan 2012). I thank an anonymous reviewer for bringing this to my attention.

3. This article focuses on descriptive representation of women. Substantive diversity and the many other identities and perspectives that comprise diversity are beyond the scope of a simple computer algorithm. I encourage scholars to be mindful of these issues when they assess, interpret, and revise their bibliographies and syllabi.

4. The tool also assesses racial diversity, corresponding to the five categories determined by the US Census using the R Package wru (Khanna and Imai 2016). However, it is far less precise and accurate than the gender estimate; for that reason, it is not the focus of this article. I urge anyone using the racial estimates from the tool to use caution in drawing conclusions from them.

5. For a good example, see this xkcd comic, available at http://xkcd.com/1425.

6. If entries always follow a set format, this is a much easier task. As a result, the tool performs best at estimating diversity for bibliographies submitted in the form of .bib files, which follow a consistent structure and eliminate the need for identifying names. For bibliographies not submitted as .bib files and for syllabi, the tool must determine which words are names, thereby leading it to be slightly less accurate.

7. For more information and a fantastic guide, see https://rpubs.com/lmullen/nlp-chapter.

8. Although both "True" and "Power" are sometimes given names, they are not common and they are more common as other parts of speech.

9. In this case, both "Power" and "True" are identified as women's names, with probabilities of belonging to a woman of 0.77 and 0.85, respectively. This introduces error into the estimate, but in repeated tests of the tool, the number of errors inserted is fairly small.

10. A shortcoming of this approach is that it assumes binary gender, and it will misgender users whose given name is not common among their gender. This is far from ideal but is, at present, largely unavoidable.

11. Of course, even if the percentages are high, scholars should be encouraged to read widely and consider a diversity of viewpoints.

12. It may be normatively desirable to overrepresent women in fields in which they are in the minorities (e.g., political methodology) on the grounds that increasing their visibility within the field may attract and retain more women and minorities who might otherwise believe that they do not belong in that field. Further discussion is beyond the scope of this article.

13. See http://womenalsoknowstuff.com.

14. See http://visionsinmethodology.org/participants.

15. See www.ruf.rice.edu/~wics/#Participants.

16. See www.saramitchell.org/journeys.html.

17. @PoCAlsoKnow is modeled after WomenAlsoKnowStuff. The creators of the account wish to remain anonymous but have asked me to mention that they want to grow and expand beyond the Twitter account to include a website and other resources. Anyone interested in helping them should tweet at @PoCAlsoKnow.

18. See http://latinocaucus.weebly.com/scholars–political-scientists.html.
19. See www.apa-politics.org/scholars.html.
20. See www.ncobps.org/search/advanced2.asp.

REFERENCES

@PoCAlsoKnow. 2016. "PoC Also Know Stuff." https://twitter.com/pocalsoknow.

Beaulieu, Emily, Amber E. Boydstun, Nadia E. Brown, Kim Yi Dionne, Andra Gillespie, Samara Klar, Yanna Krupnikov, Melissa R. Michelson, Kathleen Searles, and Christina Wolbrecht. 2016. "Women Also Know Stuff." http://womenalsoknowstuff.com.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *Quarterly Journal of Economics* 125 (3): 1101–44.

Cassese, Erin C., Angela L. Bos, and Lauren E. Duncan. 2012. "Integrating Gender into the Political Science Core Curriculum." *PS: Political Science & Politics* 45 (2): 238–43.

Colgan, Jeff. 2015. "New Evidence on Gender Bias in IR Syllabi." *Duck of Minerva*, August 27. http://duckofminerva.com/2015/08/new-evidence-on-gender-bias-in-ir-syllabi.html.

Khanna, Kabir, and Kosuke Imai. 2016. "wru: Who Are You? Bayesian Prediction of Racial Category Using Surname and Geolocation." R Package Version 0.0-2.

Maliniak, Daniel, Ryan Powers, and Barbara Walter. 2013. "The Gender Citation Gap in International Relations." *International Organization* 67 (4): 889–922.

Mansbridge, Jane. 2013. "Explaining the Gender Gap." *Monkey Cage*, September 30. www.washingtonpost.com/news/monkey-cage/wp/2013/09/30/explaining-the-gender-gap/?utm_term=.66531426550f.

Mitchell, Sara McLaughlin, Samantha Lange, and Holly Brus. 2013. "Gendered Citation Patterns in International Relations Journals." *International Studies Perspectives* 14 (4): 485–92.

Nexon, Daniel. 2013. "The Citation Gap: Results of a Self-Experiment." *Duck of Minerva*, August 16. http://duckofminerva.com/2013/08/the-citation-gap-results-of-a-self-experiment.html.

Wais, Kamil. 2015. "genderizeR: Gender Prediction Based on First Names." R package version 1.2.0.