Citations and Inequities

WORKING GROUP MEMBERS
John A. Garcia (Chair), University of Michigan
Andre Benjamin, Oklahoma University
Camille Burge, Villanova University
Carrie Carrier, Texas Christian University
Karam Dana, University of Washington-Bothell
Melanie Sayeri Dominquez (Research Assistant), University of New Mexico
John Ishiyama, University of North Texas
Ashley Jardina, Duke University
Eric Juenke, Michigan State University
Natalie Masuoka, University of California, Los Angeles
Christopher Parker, University of Washington
Stephanie Pousoulides (Research Assistant), Duke University
Gabriel Sanchez, University of New Mexico
Erica Vallejo (Research Assistant), Michigan State University

ADDITIONAL ACKNOWLEDGEMENTS
Kimberly Mealy, Betsy Super, and Erin McGrath (APSA)
# Table of Contents

**Scope and Tasks for the Working Group** ............................................................... 3

**Research, Citations, and Careers: A Meta-Analysis of The Use and Implications of Research Citations in Political Science** ........................................................................................................... 4

- What Information about Research Do Citations Provide? ........................................... 5
- Use of Citations and Issues: Assumptions ..................................................................... 6
- Citation Sources and Issues.......................................................................................... 7
- Metrics and Approaches: An Example ......................................................................... 8
- Research, Citations and Contributing Factors ............................................................... 10
- Recommended Changes to Address Citation Biases ..................................................... 27

**Journals and Editors and Citation Practices and Guidelines** ..................................... 28

- Editorial Experiences with Addressing Citation Disparities ........................................ 28
- Recommendations and Suggested Guidelines ............................................................. 34

**Race/Ethnicity and Citation Counts of APSA Members** ........................................ 36

- Characteristics of the Database .................................................................................. 36
- Source of Citation Counts ............................................................................................ 37
- The Distribution of Citations by Subgroups ................................................................. 38
- The Factors Associated with More Citations ............................................................... 39
- Conclusion .................................................................................................................... 40

**References** ................................................................................................................. 42

**Appendix A: Meta-Analysis References Regarding Research, Citations and Professional Development** .................................................................................................................. 56
Scope and Tasks for the Working Group

The role of citations in the conduct of systematic research and its impact on professional careers is often under-examined. That is, the development of citation references services, their construction, and use by departments and researchers is critical. Citations are commonplace and serve as an integral element in the development of knowledge bases and building on past research in innovative manners. Professionally, the recognition of a scholar’s research and contributions operates under a system of rewards as well operating within a variety of biases. The major focus of this APSA presidential task force is the obstacles that impede the full development of professional careers among marginalized groups of scholars. This working group’s scope lies with an analytical and evaluative consideration of aspects relative to citation research literature and its implication for all researchers.

The working group members met early in this process to discuss, comment, and think through how we could best provide recommendations to the discipline. As a result, there were five specific tasks that our working group felt it could attempt to explore in the use of citations and its impact on generating inequities for marginalized scholars. The descriptions below delineate each of these tasks.

1) Provide an overview assessment of the patterns of citations (which might include a survey of what motivates individuals’ citation choices, etc.). What becomes the basis for evaluation of the use of citations and their impact on scholars’ careers and advancement? This would include a review of the extant literature and perhaps some original databases. A key consideration may be whom to target, possibly using the APSA membership list with some stratification filters. A sub-group would take the lead in constructing items.

2) Rethink how scholars are trained in terms of citing work. For example, what do syllabi at major PhD-granting graduate departments look like? Research projects and subsequent citations often begin with what individuals read as graduate students. For example, we could conduct a coding analysis of field seminar courses (American Politics, Comparative Politics, International Relations, and Political Theory) in selected graduate programs to see the extent to which scholars who study race, sexual orientation, and gender are listed on the syllabus. The topics covered would be an important insight since these topics affect the types of scholars who are included. (i.e. inequities in scholarship on the presidency is different from that in voting behavior since the extent of diversity varies by subfields). Another suggestion is to conduct a survey on graduate training—what are the ways we teach how to write literature reviews or create syllabi. How much attention is paid to the types of scholars included? The working group on graduate education could be a collaborator on this query. A final deliverable would be guidelines on how to be more aware of inequities in citation practices.

3) Develop some guidelines for editors of our journals and to survey their current practices when it comes to addressing citation gaps. This task could be complemented with a survey of present and former editors of several more explicitly political science journals. Items include knowledge of sub-field content of submitted article and reviewer assignment, how appropriate reviewers are
selected, the role and function of citations in the review process, views on diversity of content and submitters, etc. A key question would be: Is there a role for APSA to promote best journal practices?

4) Meet the needs of those dealing with tenure and promotion and providing guidance for chairs and departments. We could create a best-practices document and/or a statement to include with external letter writer requests, which could include experiences of working group members’ experiences on the use of citation reference guides, and how they are integrated in the evaluation process. We can partner with the working group examining promotion and tenure to complement common interests. We might also explore what information can be gleaned from APSA survey of departments.

5) Conduct citation-based research to add to the knowledge base (maybe on scholars of color and scholars who identify as LGBTQ+), and work with APSA in its contract specification with an outside company. Political scientist Natalie Masuoka has forwarded the Grofman et al. data set (2017) and codebook of related information to this working group. This working group can assess additional analyses as well as augmented data. From discussions among working group members and a meta-analysis of numerous articles (see Appendix A), an understanding emerged that the challenge is that women and scholars of color tend to cluster in fields and they are less likely to be promoted at research-intense R1 institutions, which means their production rates are influenced systematically by institutional factors. These features are often missing from quantitative analyses of citation data.

From these five tasks, we decided to organize our chapter into three major sections and an appendix that includes an extensive bibliography on research, citations, and professional development. The sections are: (1) a meta-analysis of the research literature on citations, biases, and implications; (2) conversations and analysis of the role of journals and editors regarding protocols, and decision-making related to citations and research; and (3) an analysis of political scientists’ citation counts and contributing factors. There is also a final section of recommendations. While we do not provide an analysis of the training and teaching related to "passing on citations" in seminars in graduate programs, we do include a meta-analysis on development and use of syllabi which is both informative and reflects the extant literature.

**Research, Citations, and Careers: A Meta-Analysis of The Use and Implications of Research Citations in Political Science**

The effort and investment of doing meaningful research is a contested concept, but one that is essential to the academic enterprise. The evaluation and influence of research not only affect our understanding of phenomena, but also influence future directions and perspectives. Accompanying the production and dissemination of research, quantitative metrics usually accompany research assessments with operationalized instructions for their proper use and are continuously evaluated via feedback loops. Thus, we are discussing research, citations, and impacts of one’s professional development. Across disciplines, the impact of research is often
assessed quantitatively using citation counts. In turn, academics’ total citations are often used to evaluate the quality, influence, and overall contributions of researchers’ scholarly output, with implications for the distribution of resources, including a scholars’ salary and access to research funds. Citations are also used to rate the quality of the faculty and departments across different universities, and they are often part of efforts to evaluate institutional excellence at a global level. In our case, we will discuss the extant literature in this vast area (i.e., creation and utilization of citation metrics, uses of citation results that affect academicians’ professional careers, areas of biases, role of professional journals and editors, and training of research recognition and citations). When articles are highly cited, we tend to assume that they have had a significant impact on the field and that the researcher is influential. Also, the importance of citation counts is likely to increase as they become easier to compile using Google Scholar, ISI, Scopus, or Web of Science.

**What Information about Research Do Citations Provide?**

While there is disagreement on how to define quality research, there is an innate property of complex and abstract concepts that is integral in producing research. As we continue in ongoing discussions of what constitutes “good research”, each discipline has developed quantitative metrics as proxy indicators. Here are some bases from which to assess good research:

- The research must certify some kind of valued achievement.
- The concept needs to be internally complex.
- The concept has to be describable in several ways.
- The meaning of the contested concept(s) is open and evolving.
- Researchers may have their own insights about proper use of concepts, but are aware of other uses and can defend their approach.
- The concepts originated from an ideal.
- Researchers must acknowledge the different uses of a concept that compete against each other and how they will advance and improve the concept.

A more abbreviated sense of what constitutes good research is reflected by Mårtensson et al. (2016) and Aksnes et al. (2019). They reveal that “researchers’ conceptions of research quality include ideas that span from correctness, rigor, clarity, productivity, recognition, novelty, beauty, significance, autonomy, difficulty, and relevance to ethical/sustainable research.” Which of these dimensions of quality or meaningfulness is predominant in a specific assessment of research quality depends on the context of the assessment? Our focus on citations lends itself to the enhanced article visibility and the diffusion of ideas in science. We use citations as a measure of scientific influence and attention. While citations are not necessarily a signal of inherent
scientific quality, citations are a metric signaling which articles receive attention, prominence, and usage in scientific fields (see the comprehensive review by Bornmann 2008). Also, the importance of citation counts is likely to increase as they become easier to compile using Google Scholar, ISI, Scopus, Web of Science, etc. Researchers use the Web of Science, Scopus, and Google Scholar (among others) to initiate new research plans and help frame the questions they want to answer; and then to search for and discover the key literature that helps to support and inform their current research. We now move our discussion to the bases and use of citation-related research and underlying assumptions.

Use of Citations and Issues: Assumptions

In this section, we discuss research that attempts to evaluate or rank political science departments by recording the number of citations their respective faculty received over a specified period. This type of citation research examines the number of articles published by faculty, the number of citations noted for each faculty, the impact factor of the respective journals, and the appearance of work in top publications, both in the aggregate and on a per-faculty basis (Peress 2019).

Peress (2019) calculated the number of citations of articles and books across a five-year period as a measure of research productivity and quality assessment. Some issues in producing accurate citation “assignment” to the right author included: individuals who had changed their names; special punctuation in their name; having a very common name; and being prolific outside the field of political science, which could produce relatively larger differences. In addition, Peress also calculated the sum of the five-year impact factors of the journals in which an author’s publications appeared. He indicated that journal publications are blind to the identity of authors to a greater degree than citations; yet it is possible that individuals teaching at prestigious universities may have their articles or books cited simply because they teach at prestigious universities. A measure based on impact factors may be less sensitive to this. The fourth measure Peress employed was based on the number of top publications published by each faculty member (Peress, 2019).

Some other notable patterns uncovered by Peress’s research found that departments with significant theory groups and significant numbers of qualitative scholars do better based on citation rankings. In addition, departments with large numbers of Americanists and quantitative scholars do better based on impact and top publications. Male and female scholars seem to agree that peer-reviewed journal articles are the main currency for stature in the realm of academia. Determining which of these measures is the most appropriate inherently entails a value judgment and is not neutral to the expectations of the different fields of political science. While Peress did produce a list of top-ranked departments, our purpose is to highlight a common use of citations to assess departments and individual faculty. The way citation counts are constructed and applied deal with the identification and use of metrics and possible limitations and biases. What have been the protocols for citation-based researchers and evaluation units?
Citation Sources and Issues

Bramer et al. (2017) concluded, “Optimal searches in systematic reviews should search at least Embase, MEDLINE, Web of Science, and Google Scholar as a minimum requirement to guarantee adequate and efficient coverage.” Especially for the natural sciences, the Web of Science is the primary source of publication and citation data for many systematic research reviews across a broad range of disciplines and about twice as many research management and evaluation studies as any other source.\(^1\) Our citations working group has also looked at SCOPUS, and there are many other citation databases, as discussed below.

Recently, researchers have directed their attention to Google Scholar (GS) as the database for citations. Scholars create home pages to record their publications, and departments and researchers use Google Scholar to note the extent of citations. The strengths of GS include incentives for quality, visibility, and open access; provision of a platform for networking; and the reliability and consistency of its citation counts. GS has biases, however, because citation is a social and political process that puts certain groups at a disadvantage, such as women and younger scholars (Kousha and Thelwall 2017; Falgas et al. 2008). GS counts also reflect practices of strategic citation that exacerbate existing hierarchies and inequalities.

GS counts can have a double effect because of its reliance on citation sites can figure into both departmental and external evaluations. Instead, the unique role and obligation of expert external reviewers is to offer a reasoned, rigorous, and qualitative assessment of a scholar’s contributions Jensenius et al. 2018). Another disadvantage of Google Scholar based upon recent studies (Jensenius et al. 2018) is that these common-sense citation practices put pioneering scholars with bold ideas, early-career scholars, scholars in smaller research communities, women, and solo authors at a disadvantage. Also, many people cite works for strategic rather than principled reasons. As a result, citation counts do not offer an accurate assessment of scholarly impact and may exacerbate existing social hierarchies and inequalities.

GS counts are biased toward incremental work and away from boldness and innovation. In general, the number of citations that an article or book receives in the five years or so after publication says little about its long-term effect. With this pattern, it is better to publish incremental work on topics in which there is a large, active subgroup of scholars who cite one another than it is to open a new field of research (Jensenius et al. 2018). Short-term, citation-centered evaluations discourage boldness and innovation, especially among early-career political scientists. Scholars in larger research communities have an advantage over scholars in smaller fields of study because they have a larger pool of people who can cite them. Regarding gender, Maliniak, Powers, and Walter (2013) analyzed over two decades of publications in International Relations (IR) and found that, controlling for a variety of factors such as publication venue,

\(^1\) One important distinction is the nature of the citation in terms of its focus. Bramer, et. al. (2017) differentiates between informetric reviews, meta-analysis, and research publications. Informetric pieces include titles or abstracts that contain keywords such as ‘scientometric’, ‘bibliometric’, ‘informetric’); while reviews contain titles or abstracts with contain phrases such as ‘systematic review. Meta-analyses include ‘literature search’. These distinctions reveal the differentiation between scientometric topics and the rest of the research literature.
methodology, and tenure status, an article written by a woman receives 80 percent as many citations as a similar article written by a man. Women are less likely to be cited by the most influential articles, and less likely to self-cite, possibly because women experience penalties for self-promotion (Moss-Racusin and Rudman 2010). As the number of women scholars increases, work by women tends to get cited more by both men and women authors. Dion, Sumner, and Mitchell (2018) concluded that “citation practices are influenced by gender diversity.”

Additional factors affecting citation counts include co-authorship and strategic citation as bases for citing specific works. People who tend to co-author can generate citations more easily than people who tend to solo author. For example, Teele and Thelen (2017) demonstrate that all-male teams authored most of the collaborative work in 10 of the most prominent political science journals. In other cases, some books and articles are cited almost entirely for their flaws, not their importance. Authors often cite poorly executed studies, easy targets, and straw-man pieces to explain what they are arguing against (Nexon and Jackson 2015). Strategic citation leads people to cite authors of works already deemed important, not because of their relevance, but because they are perceived as gatekeepers, hold key editorial positions, or reside in powerful departments.

GS does have advantages: It promotes consistency in research evaluation, encourages transparency, publicity, and openness, makes it easier to gain access to scholarly work, facilitates networking among scholars, and may provide incentives for quality over quantity. On the other hand, GS citation counts favor incremental work, scholars in larger research communities, male—and likely white—scholars, scholars who co-author, and work that is cited strategically. The uncritical use of GS entrenches long-existing inequalities in the political science discipline. As noted in our discussion of the use of citation services, biases do exist. We will continue our discussion of citation metrics and approaches to conducting citation-related research.

**Metrics and Approaches: An Example**

Publication- and citation-based metrics can be weighted based on the type of publication (journal article, book chapter, monograph, etc.), a quality rating of the publication venue (which can be decided by committee or taken from a trusted source such as the ISI Web of Science by Thomson Reuters), and the number of authors. Different weighting schemes are in use, and much has been written about how to implement them fairly, such as the work by Aagaard et al. (2015) and Piro et al. (2013). It is plausible for publication and citation-based metrics to be positively correlated with other metrics targeting research quality, and indeed this correlation has been shown to exist, for instance, by Jarwal et al. (2009).

In a study by Leydesdorff et al. (2011), data are drawn from the Web of Science Core Collection using the Science Citation Index Expanded, the Social Science Citation Index and the Arts and Humanities Citation Index (SCIE, SSCI and AHCI) for the five-year period of 2014–2018. Documents for analysis were restricted to original academic journal contributions (i.e., articles and reviews). Because citation counts do grow over time at rates that are field-dependent,
Garfield (1979) calculated Category Normalized Citation Impact (CNCI) for each individual paper.

Clarivate’s Distinct Author Identification System (DAIS) uses a weighted comparison of author clusters drawing on over 20 points of distance/similarity from publication metadata, including the author’s ORCID, name, subject category, use of references, author-based co-citation analysis (Small 1973; White and Griffith 1981), institutional name, etc. The concept of citation clusters allowed the researchers to construct article-level clusters based on either direct citations or co-citations and reconciled them with major and minor subject categories in the All-Science Journal Classification. Salton’s (1979) analysis was in the computer science field, in which three main area were identified: (1) theoretical foundations, such as theory of computation, (2) hardware and computer systems, such as architecture, and (3) software, such as programming systems. Other stipulations to define the universe of publications included that each publication had a digital object identifier (DOI) and could also be matched to article identifiers in the Scopus bibliography. This action was justified since crossmatching DBLP publications to records in the Scopus abstract and citation database of peer-reviewed literature enabled them to harvest the richer links in Scopus, as well as extract links to publications from other disciplines.

Salton used direct citation links as the basis for cluster formation, and co-citation to obtain an alternative view. In applying both clustering by direct citation and by co-citation, he attempted to consider, wherever possible, the criteria articulated by Šubelj et al. (2016) that (1) the largest cluster should be no more than 10 times the smallest one; (2) small clusters should be eliminated; (3) small changes and replicates should yield similar results (“stability”); (4) computing time should be minimized, where possible; and (5) the clustering should seem reasonable on a qualitative level (“intuitive sensibility”).

The concept of clustering and conductance may require some additional clarification. Clustering is often described as that of finding a set (or sets) of vertices that are more related, in some sense, to each other than to other vertices in the graph (Shun et al., 2016). For example, Andersen and Lang (2006) use a variant of the algorithm of Spielman and Teng (2004) to identify communities in networks. Leskovec et al. (2009, 2010) and Jeub et al. (2015) use the algorithm of Andersen et al. (2006) and other graph-clustering algorithms to study the properties of clusters of different sizes in social and web graphs. The conductance of a cluster is a widely used metric to measure cluster quality. Intuitively, low-conductance vertex sets tend to correspond to higher-quality clusters because these sets are larger and have fewer edges to vertices outside of the set. Analysis by Shun et al. (2016) evaluated clustering by direct citation (smaller is better), noting that conductance has been found to be a good metric for this purpose (Emmons et al. 2016; Almeida et al. 2012). In each case, the highest-numbered cluster had the greatest conductance value and the smallest number of nodes.

Shun et al. (2016) used co-citation as an alternate view of these DBLP data (clusters using co-citation, the frequency with which a pair of articles is cited by other articles (Small 1973; Marshakova-Shaikevich 1973). Co-citation, first described independently by Small (1973) and (Marshakova-Shaikevich 1973), provides insight into the emergence of new ideas derived from
the association of previously independent ones. Unlike clustering by direct citation, where every input publication is assigned to a cluster and every citation is weighted equally, the co-citation relationship between papers is weighted to represent the strength of the co-citation history.

These data suggests that the fields central to computer science in the 1970s (Salton and Bergmark 1979) were a historical triad of hardware, software, and theory which are more likely to be found in multiple clusters than peripheral fields. A second inference is that, in some cases, journal-based classification and their article clusters align well (i.e., hardware and architecture). A third inference is that the ASJC minor subject area “Computer Applications” is relatively broad, and publications thus labeled are present at the ≥ 15% level in 16 out of 20 clusters. Finally, for this DBLP dataset, as we clustered it, interactions with fields outside computer science such as biology (i.e., biochemistry and neurology) are detected in two separate clusters. The field of computer science has not only expanded in volume, it has expanded in its interactions with other fields, and has also resulted in new disciplinary and interdisciplinary subfields.

These works on citations, clusters, and networks in computer science have relevance as we examine political science. For one, the number of fields and sub-fields do represent both a distribution of research scholars among many content areas with varying sized communities. Secondly, networks exist across scholars for both the production of research efforts and recognition of specific research products. As a result, clustering and interactive networks affect the extent of citations generated by these clusters. Finally, as with many fields in the natural and social sciences, there are cross-disciplinary interactions in which limiting citations within an established discipline provides a limited view of research activity and impact. We now move to some detailed discussions of elements or factors that affect the use and numbers of citations attributed to researchers.

Research, Citations and Contributing Factors

Gender

There is a growing literature analyzing “gender effects” related to citation counts and academic recognition. For the most part, women researchers are cited less than their male counterparts. There are, however, some studies that find no gender differences, and a few find an advantage for women. The remainder of this section will present extant findings and explanations.

Mitchell and Hesli (2013) used NSF data to show declining percentages of women in the discipline as ranks increase, noting that women constitute 40% of doctoral degrees in the field but only 28% of APSA members in 2009. A decade later, it was still only 33.6% of members (APSA 2018). These data accompany other estimates of women’s participation in professional associations (Breuning and Sanders 2007). Similarly, Teel and Thelen (2017) noted that women comprise 27% of faculty in the 20 largest PhD-granting departments, 31% of APSA members, and 40% of PhDs in political science. Hancock, Baum, and Breuning (2013, 6) reported that among International Studies Association members, 20% of women are full professors, compared
with 34% of men. These types of aggregate disciplinary snapshots identify the population of female scholars in our profession. They do not, however, identify nuanced differences across disciplinary subfields or narrow substantive areas of interest, which often have significant variations in gender distributions. In 2018, significantly more women identified their primary research or teaching field as public policy (41.4%) or comparative politics (39.3%) than the overall female representation in APSA (35.8%). In contrast, women are significantly underrepresented among members who claim political philosophy and theory (31.6% female) or political methodology (19.4% female).

Dion et al. (2018) found no evidence that the low percentage of female authors simply mirrors an overall low share of women in the profession. Instead, they found continued underrepresentation of women in many of the discipline’s top journals. They also found that women are not benefiting equally in a broad trend across the discipline toward co-authorship. Most of the published collaborative research in these journals emerges from all-male teams. They also found that women submit at lower rates than men, but fare comparatively well in the review process, appearing in the journals at somewhat higher rates than their presence among all submitting authors. Dion et al. (2018) think that differential rates of submission by women may well be a problem beyond the limited journals and time periods these authors were able to examine, and we are partnering with several journal editors to assemble the data needed to explore this issue. Teele and Thelen (2017) also found evidence that women are not participating equally in a broad move across the discipline toward co-authorship. In addition, it appears that the methodological proclivities of the top journals do not fully reflect the kind of work that female scholars are more likely than men to publish in these journals. The underrepresentation of qualitative work in many journals is associated as well with an underrepresentation of female authors. Women publish less quantitative research than their male counterparts, and the top journals are publishing mostly quantitative research.

Another aspect of gender differences and citations is that they do not identify nuanced differences across disciplinary subfields or narrower substantive areas of interest, which often have significant variations in gender distributions. Comparisons of organizational membership and published authors also reveal potential gendered publication gaps if women’s representation as article authors is significantly less than their presence in a field. There continues to be a pronounced underrepresentation of women in many of the top journals in political science.

Teele et al. (2017) saw signs of gendered patterns of exclusion from Comparative Political Studies, International Organization, Journal of Conflict Research, Journal of Politics, Political Theory, and World Politics. The data (Teele et al. 2017) show that women are overrepresented among the authors of published qualitative studies, and they are underrepresented in the pages of those journals that publish very little of such research. The two journals that publish the fewest qualitative articles (America Journal of Political Science and Journal of Political Science) also publish the fewest female authors. Thus, women are underrepresented in work across all methods except qualitative work and political theory. The gap is most severe in work that relies primarily on formal theory, followed by conceptual and experimental work. By contrast, among articles
categorized as employing primarily qualitative methods (case studies, small n), the share of women exceeds their membership share in the APSA.

Rossiter (1993) coined the term “the Matilda Effect” to describe cumulative disadvantages attributable to gender inequality. Given the historical and continued underrepresentation of women in the field, especially in its upper echelons, economics also provides a vivid case study to examine the relationship between gender, visibility of research, and cumulative advantage in science. Potential exists for “old boys’ clubs” in both “upmarket and down-market journals” in the discipline. Status characteristics and social capital associated with gendered networks provide advantages in access to academic, professional, and social information, which facilitate publishing opportunities.

Some of the contributing factors affecting this citation gender gap are:

1. Men and women tend to work at different institutions. The 2006 TRIP of IR scholars finds that men are more likely than women to be employed by PhD-granting institutions, while women are more likely to be employed by liberal arts schools.

2. Women may publish less in the early years of their careers because of their need to take parental leave. Additionally, works have found that new fathers in the discipline are more likely to use parental leave to advance their career through publishing, while women are more likely to use the leave to focus on childcare responsibilities (Antecol, Bedard and Stearns, 2018). This might not affect productivity over the long term, but if citations depend in part on building name recognition, then fewer publications early in one’s career could translate into fewer citations over time. Even in fields such as women and politics, in which female scholars comprise most of all authors, male authors in Politics & Gender are still 14% less likely than female authors to cite the work of women (Dion, Sumner, and Mitchell 2018).

3. Men and women tend to study different substantive issues. Men are more likely to write articles on security, US foreign policy, and methods. Women are more likely to write articles on human rights, comparative foreign policy, health, international law, and the environment. If these topics are less popular and less well cited, then this could also help account for the gender gap in total citation counts.

4. Men and women report using different theoretical paradigms to analyze international politics. Women are more likely than men to publish articles that are constructivist or non-paradigmatic, while men are more likely to publish articles that are a theoretic, realist, or Marxist.

5. As the 2006 TRIP survey and Breuning and colleagues show, men and women tend to situate their work in different epistemological schools. Men are slightly more likely to report that their work is positivist. Women are nearly twice as likely as men to report their work being postpositivist.
6. The TRIP surveys show that women are just slightly more likely to report employing qualitative methods than are men. Conversely, men are only slightly more likely to report employing quantitative methods than women. Men are also much more likely than women to report using formal methods.

7. It could be that the citation gap is due to the venue in which men and women publish. Women may tend to publish in certain journals, and it is these journals that tend to draw fewer citations than others.

Maliniak (2008, 2011) reveals that a subset of scholarship in IR is being systematically undervalued, even if inadvertently. Research produced by a woman will be read less and cited less than research produced by a man. Not only does this mean that the trajectory of intellectual developments will be slower than it should be, but it means that the types of topics and methods being showcased in journals and on syllabi are likely to be skewed toward those favored and pursued by men. When IR scholars tend to cite along gender lines, then any researcher who is male will get a boost in citation counts simply for being a man. Several processes could produce publication gaps, including (1) the leaky pipeline, or fewer women in senior ranks; (2) lower article submission rates of women compared to men (Djupe, Smith, and Sokhey 2019; Hesli and Lee 2011); (3) the rise of co-authorship, which benefits primarily male authors (Teele and Thelen 2017); and (4) gender biases in editorial decision-making processes.

An extensive study by Maliniak et al. (2013) develops multiple models to explain possible gender differences by accounting for a wide range of other factors. For example, their models accounted for controls of paper age and the square of paper age, reflecting the expectation that older papers will receive more citations because older papers have had more opportunity to be cited. Other controls included R1 author’s status, tenured author status, co-authored papers, and the journal that published the paper, reflecting the expectation that potential citers are more likely to be more aware of papers published by researchers at R1 institutions, published by tenured researchers, published by multiple researchers, and published in more familiar journals. The models developed by Maliniak et al. (2013) contained dichotomous controls for the type of research reported: positivist research, research with ideational factors, research with material variables, research in particular issue areas, and research with particular methodological approaches.

Some of their findings included the interaction of journal familiarity score and the female authorship variable was substantively and statistically significant (0.47, \(p=0.013\)). Removing the journal familiarity score in their model No. 8 and restricting the analysis to the three most familiar journals (APSR, AJPS and JOP) returned a point estimate for the gender citation gap of 0.58 (\(p=0.012\), \(n=165\)) compared to the gap of 0.95 for the nine residual journals (\(p=0.562\), \(n=1367\)), providing corroborating evidence that the gender citation gap in international relations is primarily an elite phenomenon. Their results generally support the inference that compared to citations of papers authored by only men, citations of papers authored by only women tend to come from papers that receive fewer citations and are thus less central to the international relations literature.
Gender difference in citation counts appears restricted to the right “distribution” tail of citations and to the most familiar journals, so any advantage for male researchers and any disadvantage for female researchers might not be evenly distributed among male and female researchers. Restricting the analysis to the right tail of citations and the most familiar journals, it is unclear whether the estimated gender citation gap is driven by bias or by unmodelled factors, given that the models contain no controls that can explain why some papers land in the right tail of the citation distribution or are placed in the most familiar journals. Many variables—including year of publication, venue of publication, substantive focus, theoretical perspective, methodology, tenure status, and institutional affiliation—produced significant results in a variety of modeling choices (Maliniak et al., 2018). This was likely because women tend to cite themselves less than men, and men—who make up a disproportionate share of IR scholars—tend to cite men more than women.

Maliniak et al. (2018) find that articles published by men are less likely to cite work by women than articles published by women. On the other hand, Østby and colleagues find that gender is not a significant determinant of publication in their analysis of submission and publication rates at the Journal of Peace Research between 1983 and 2010. Using these data, Malinak et al. (2018) demonstrate the existence of a persistent gender gap in citation counts: Articles written by women are consistently cited less than articles written by men. Women are also more concentrated on the periphery of the IR network, where their work is cited less often by authors of the most heavily cited work. These findings may explain the underrepresentation of female authors in syllabi (Colgan 2017; Hardt et al. 2017), edited volumes (Mathews and Andersen 2001), and textbooks (Cassesse, Bos, and Duncan 2012). This is especially true of women who have not yet earned tenure. Observable differences between male and female IR scholars—including productivity, institutional affiliation, publication venue, or epistemology—cannot account for this gap.

The most common form of publication across all the journals is a single male author (41.1%), while the second most common form of publication is an all-male team (24%). Women working alone represent 17.1% of publications, while all-female teams represent just 2.4% of all journal articles (Teele & Thelen 2017, Mitchel et al. 2013, Breuning & Sanders 2007). Finally, cross-gender collaborations account for 15.4% of publications. The norms of co-authorship have changed over time and differentially across genders. Fisher and others (1998) show in their analysis of co-authorship in three political science journals that while co-authorship overall across these journals has increased, it has increased more quickly for women. The modal collaborative team is either all male or mixed gender—never all female. Four journals (Journal of Race, Ethnicity, and Politics; Journal of Experimental Political Science; Public Opinion Quarterly; and Political Communication) have more mixed gender teams than other types of author configurations. Five journals (American Journal of Political Science, Political Analysis, Journal of Conflict Resolution, British Journal of Political Science, and Journal of Politics) have mostly male-only, collaborative author teams. This reflects tendencies for women to engage in fewer collaborative publications and to work in fields (e.g., comparative politics) in which collaboration is less common. For example, Politics, Groups, and Identities’ submission report...
indicated the number of co-authored, women-only manuscripts is lower (3 or 9.1%) compared to co-authored, male-only manuscripts (5 or 15.2%).

Symonds and colleagues (2006) found that in cross-gender collaborations, women are nearly four times more likely to collaborate with men than men are to collaborate with women. The authors document an increase in multi-authored manuscripts over this period and show that by the mid-1990s almost half the articles that appeared in these three journals were co- or multi-authored. Women were nearly four times more likely to publish articles co-authored with men than the other way around (1998, 852). Articles published by exclusively male multi-author groups rose as a percentage of all articles published across all three journals to 39.2%, while cross-gender collaborations rose but only to 13.5%.

The underrepresentation of women in multi-authored work appears not to be function of their unwillingness to engage in collaboration. Symonds et al. (2006) show that on average, 51% of publications with at least one female author emerge from collaborative work, while another table shows that 47.4% of publications with at least one male author are team projects. But while most of the work with at least one male author emerges from all-male teams (28.8%), only 6.6% of the publications with at least one female author are by all-female teams. The number of articles published by female authors working alone has been mostly constant, which means that the (relatively small) overall increase in the share of female authors comes from co-authorship with men.

Whether single or multiple authorship, the penchant to self-cite is another means to enhance citation counts. Self-citation may reflect some strategic decisions to use this opportunity to promote greater visibility and increased citations. This practice indicates fewer self-citations by women, and more within gender citations. Second, self-promotion strongly affects citation counts, and women are less likely to promote themselves (King et al. 2018). Not only does self-citation increase one’s overall citation score, but it also exposes one’s work to a larger number of scholars, exponentially increasing citations. The fact that citation counts can be manipulated, and that this strategy is more apt to be pursued by men, disadvantages women. Scholars tend to cite work by scholars of the same gender. This does not necessarily pose a problem for citation counts if a field has a fairly equal number of male and female scholars. If, however, a field is heavily dominated by men, as is the case in IR, then this pattern will lead to significantly fewer citations for women and significantly less exposure for their scholarship. It also means that the gap is not likely to disappear until a more equal number of male and female researchers exists. Our previous discussions reinforce that networks matter. Producing high-quality work is not sufficient for research to gain the attention of the widest number of scholars or have the greatest impact. Scholars tend to cite scholars they know, and work produced by lesser-known scholars, or scholars in small or peripheral networks, is likely to be cited less. If networks tend to bifurcate along gender lines, then any field that is disproportionately male will also disproportionately favor their work.

Several processes could produce gender publication gaps, including (1) the leaky pipeline, or fewer women in senior ranks; (2) lower article submission rates of women compared to men...
(Djupe, Smith, and Sokhey 2019; Hesli and Lee 2011); (3) the rise of co-authorship, which benefits primarily male authors (Teele and Thelen 2017); and (4) gender biases in editorial decision-making processes. Building more information about women’s representation in specific areas of research makes it easier to know whether the distribution of cited authors is biased, even when calculating the gender and racial breakdown of references (Sumner 2018). She argues that scholars should consider gender representation in their research areas if they want to minimize implicit biases in their citation practices.

The fact that observations and documentation measured by Teele and Thelen (2017), show significant variation across journals in publication rates for women would suggest that we should be skeptical of explanations of the problem that “attribute it to its victims”. They intimate that the APSA sections that sponsor journals should evaluate whether the publications provide ample descriptive representation of section members. In addition, those that select journal editorial teams should pay attention not only to their diversity, but also to their plans for addressing potential citation biases. In addition, they should consider using tools like the Gender Balance Assessment Tool (Sumner 2018), which can help political scientists quickly and easily evaluate gender balance in their bibliographies. Expressed somewhat differently, gendered publication or citation gaps between membership and authorship in related academic journals can provide insight into research areas where potential biases in pipelines, co-authorship, and submission rates remain substantial. Data by Dion et al. (2018) provides more nuanced information about relevant gendered baselines for scholars who wonder whether they are missing research by women in their articles, books, and syllabi, as well as those who want to identify research areas in which gendered biases in publications and citations may be most significant.

Parenthood, Citations, and Research

Extending the focus between gender citation gaps, a concomitant factor is the effects of parenthood on the publication output of male and female academics. Lutter and Schröderr’s study (2015, 2016) at German universities indicated that having children leads to a significant decline in the number of publications by women on average, while not affecting the number of publications by men. Women publish about 20% less than men after controlling for the adverse effects of children on productivity. The gendered effect of childbearing depends partly on prior levels of women’s academic achievements, suggesting a mechanism of performance-driven self-selection. Lower-performing women tend to suffer a stronger motherhood penalty than better-performing women, while the publication output of successful women (who have been granted academic awards) is not reduced through childbirth (Lutter 2015).

There is growing research literature supporting the hypotheses that confirm that motherhood decreases the publication output of women (Hargens, McCann, and Reskin, 1978: p. 161; Long, 1990); yet others argue that only young children decrease publication output (Kyvik, 1990: p. 158; Kyvik and Teigen, 1996: p. 69; Stack, 2004: p. 914; Mason, Wolfinger, and Goulden, 2013: p. 29). Some do not find any effect. Hamovitch and Morgenstern (1977: p. 643) were the first to conclude that, after controlling for academic rank, teaching duties, years since PhD completion, type of university, and academic discipline, “the remaining sex difference in publication (20 per
cent fewer articles published by women than by men) is not due to any relationship or tradeoff between children and productivity.” At the same time, they suggest that there might be a motherhood penalty for very young mothers or those with two or more children. They also find that fathers increase their research productivity after their first child (Lutter and Schroder, 2019). Most studies have been based solely on cross-sectional or between-level effects. Criticism of these studies point out that they compare groups but do not consider individual career changes (i.e., within-level effects, obtained through fixed-effects regressions). Suggestions to control for possible selection effects before childbirth, as researchers who have been assured of their academic potential early on may tenaciously stick to an academic career and continue to publish even after having children (Cole and Zuckerman, 1987: p. 125; Fox, 2005: p. 145; Joecks, Pull and Backes- Gellner, 2014: p. 520). In line with Joecks, Pull and Backes-Gellner (2014), children may influence publication output through (1) effects of self-selection, (2) effects of time constraints, and (3) incentive effects as well as changes in preferences. The first mechanism, self-selection, may operate in two opposed ways, depending on career orientation and how much academic acclaim researchers received.

Co-authorship

While co-authorship has always existed in research, some current trends indicate a rise in co- and multi-authorship. How does this affect citation counts among members of collaborative works? Are there gender and racial-ethnic patterns for such group endeavors? More recently, studies of bibliographic data suggest a strong correlation between the growth of citation networks and their corresponding co-authorship networks. By tracing the change in citations they exchanged and, by tracing the shortest path between authors in the co-authorship network, Singh (2014) has used probability methods to quantify the correlation between citations and shortest paths, and the effect on the dynamics of the citation-co-authorship system. Singh found that author pairs who have a co-authorship distance $d \leq 3$ significantly affect each other’s citations, but that this effect falls off rapidly for longer distances in the co-authorship network (Singh, 2014). Moreover, the citations received by most authors are due to reciprocal citations from current or past co-authors.

An awareness to the state of the art and motivation to solve open problems become factors in setting up new collaborations between individuals. Researchers attend gatherings and conferences to broaden their scope of a subject area and look for new ideas and possible collaborations (Singh, 2014). As we indicated, the pattern of collaboration has shown a shift from individual efforts to more cooperative research, increasing the productivity and diversity of scientific publications globally, and resulting in an increase of innovation in this century (Dong et al. 2017). Every new publication result in co-authorship and citation events, which suggests that citation and co-authorship networks are interrelated and should have a strong positive correlation. Many studies addressed these networks and pointed out strong interdependent relations between evolving citations and co-authorship networks (Kas et al. 2012; Keegan et al. 2013; Martin et al. 2013; Amblard et al. 2011; Ding 2011; Tol 2011; Glänzel and Thijs 2004). Combined co-authorship and citation networks have been used to predict new collaboration opportunities (Lande and Andrushchenko 2016).
Post-1993, Singh (2014) noticed a sudden increase in the number of such pairs. This sudden change is because of the introduction of papers with a high number of authors in that period. These papers lead to large cliques in the co-authorship network. Most of these papers are published by large collaboration groups often having multiple common authors in their publications and with many Indian authors being part of such groups. Also, Singh observed high citation reciprocity at initial times. As the network grows, the distribution of citations becomes more heterogeneous as some authors receive more citations than others (authors of influential papers receive many citations). Authors that are no longer publishing cannot reciprocate anymore but still receive citations. Thus, more citations are given out than received on average, which results in a decreasing trend in citation reciprocity.

Singh calculated the average fraction of incoming citations that an author responds to by subsequently citing the author who initially cited her. When the co-authorship network is in its initial phase, with a small number of researchers, most co-author pairs cite each other. So, a high citation reciprocity exists initially. As the network grows, the distribution of citations becomes more heterogeneous as some authors receive more citations than others. Such projects are likely to produce papers that cite other publications related to the same project and to have a high number of overlapping authors. This leads to the observed large number of self-citations. While all authors do cite others in their papers, few of them receive citations. Therefore, while the outgoing self-citations increase rapidly, the incoming self-citations observe a dip.

One reason is that there is an increase in the heterogeneity of references because of increasing research complexity and interdisciplinarity. Another possible explanation is that improvements in bibliographic technology resulting from online research resources (Web of Science, Scopus, Google Scholar, and the like) and increasing online journal access have made it easier for researchers to find and cite large numbers of articles. A third potential explanation is that the increasing number of references is a function of the increase in the number of publications—authors cite more articles simply because there are more articles to be cited. We find that co-authors dominate the citation patterns in our networks with a heavy bias (≈ 75%) toward self-citations (Singh 2014). The remainder of the citations were mostly between pairs that have a short (d ≤ 3) connected path in the co-authorship network; the average number of citations exchanged decays with increasing co-authorship distance.

An even more recent development with multiple authorships is the number of contributors. There has been a general increase associated with multi-authorship (more than 10 authors and more than five countries); and even a more perturbing outcome of hyper-authorship (more than 100 authors spread across more than 30 countries (Adams, 2013).) Citation impact increases with rising authorship but gains at higher counts are more evident and variable for smaller countries (Adams 2013). Although multi-authorship leads to higher impact, this link is coherent, progressive, and regular for most (but not all) discipline categories, and for some fields there is little or no effect (Adams 2013). At the same time, there are some concerns (Cronin 2001) about what a name in an author list means in contemporary terms: Is an author still synonymous with a writer? There are also questions of where accountability and collective responsibility lie as the authorship of a paper rises (Croll 1984). Is a paper with 100 authors the same kind of
communication as a paper with one or a few authors? Research with that many authors is concentrated in physics, including work at international space observatories and at CERN. Cronin (2001) concluded that hyper-authorship signifies a change in research nature. The major challenges of research—population studies, epidemiology, climate change, particle, and space sciences—require investment in equipment, data collection, longitudinal studies, and analytical processing associated with large teams. How this pattern of large numbers of authors affects the social sciences is a possible dimension of citation counts and status recognition.

Citation of Books

Knowing the relative stature of journals and presses helps scholars make professional decisions and can help colleagues, department chairs, and administrators make difficult personnel decisions and evaluate individual members and entire departments (Samuels 2013). The average political science book published by a university press receives about three times the number of citations received by an article indexed in the SSCI. Books’ influence varies by subfield, with books published in methodology receiving many more citations on average than books published in other subfields, followed by books in International Relations. Overall, books published on American politics are cited least frequently (Samuels 2013).

Samuels attempts to quantify the scholarly influence of published books, counting citations after five years of all political science books published in 2004 and 2005. This effort offers a reality check on political scientists’ subjective perceptions of publishers’ relative quality (Samuels 2013). Counting the citations that books receive in articles as well as in other books provides a comprehensive assessment of book influence and publisher quality and provides a benchmark for judging any book’s relative importance. Books published in international relations tend to be cited relatively more frequently than books in the other substantive subfields, even considering only citations in English-language publications. Yet the biggest revelation is that books on social-science methodology receive at least twice as many citations on average as books in any substantive subfields (Samuels 2013). Adding up total citations, the average is about 23.4 citations per book after five years. However, this number obscures a substantial difference between university presses (34 presses, N =1,517) and others presses (37 presses, N =1,666). The average book published by a university press is cited almost 2.5 times as often (33.2 citations) as the average book published by other presses (14.4 citations). This means that the average peer-reviewed article indexed in the SSCI receives about one third the total number of citations as the average university-press book after five years (Samuels, 2013).

The placement of which publisher prints the book is debated so that the only thing that should matter is whether one’s research demonstrably shapes the scholarly conversation. To the extent one accepts that dictum, a subjective assessment of publisher quality is inferior to a citation count. Although not surprisingly, publishing with a university press tends to be a safer bet (Samuels 2011,2013). At the same time, publishing with a university press is more likely on average to influence the scholarly conversation, as the average number of citations for a university-press book is about 2.5 times that of a book published by other presses. Books published in American politics follow the same pattern as articles in American politics (Samuels
They are least-frequently cited relative to books published in other subfields, particularly because books in that subfield are relatively less frequently cited in peer reviewed articles. Finally, there is a site on Google called Google Books (https://books.google.com/), which is distinct from Google Scholar, that does not under-count citations of books that appear in other books (see Samuels 2011). This makes it possible to ascertain the number of times a book is later cited in *other books*. The protocol usually entails entering the book’s full title with these search parameters: (1) search for “all books”; (2) limit search results to “books” (eliminating “magazines” from the results); and (3) limit results to books in English; and (4) limit search results by date. This section on book citation was added to provide a fuller scope of research in addition to articles and ways to interpret citation counts.

**Citations and Syllabi**

Another area in which citation patterns are developed occurs during one’s graduate education. That is, graduate courses and readings are driven largely by the faculty’s choice of readings and research publications. The selection of included readings can direct a graduate student to recognize good research and sources for such research. There have been studies that examine course syllabi and the choices made by faculty. Smith et al. (2020) analyzed the Graduate Assignments Dataset (GRADS), which is the most comprehensive dataset of assigned graduate readings to date. GRADS includes 75,601 syllabi readings from 840 syllabi and 605 unique instructors at 94 US-based political science departments.

The report noted several findings. First, overall, instructors infrequently assign female-authored scholarship relative to the rates at which women publish. Second, instructors who are women, people of color, and those from more gender-equal countries assign significantly more female-authored readings than white male instructors and those from less gender-equal countries. Third, among women, instructors assign more female-authored work. We suggest that women’s underrepresentation on syllabi may contribute to “the leaky pipeline,” which describes women’s attrition from academic careers (Smith et al. 2020).

This underrepresentation is particularly pronounced when considering women as first authors. Instructors’ characteristics affect whether they assign work by women; both identity and socialization also appear to play a role. Among female instructors, generational cohorts also matter. Many faculty tend to assign some of the same readings that they themselves read as graduate students and to seek out relevant syllabi from other instructors. They also may rely on classic works and “elite readings,” in which gender gaps in citation counts are largest (Zigerell 2015). Instructors are likely to assign work by well-known scholars; male authors, in general, are likely to occupy more central locations in scholarly networks (Smith et al. 2020). Implicit gender biases could lead instructors unconsciously to favor male-authored readings. Beginning at a young age, individuals adopt the gender stereotypes that women are less brilliant than men and less capable academics (Bian, Leslie, and Cimpian 2017; Leslie et al. 2015; Williams, Phillips, and Hall 2014). Scholars have found evidence of gender bias in academia in evaluations of scholarly work, letters of recommendation, and certain hiring practices (Knobloch-Westerwick and Glynn 2013; Krawczyk and Smyk 2016; Vander Lee and Ellemers 2015; Madera, Hebl, and
Martin 2009; Rivera 2017). For example, recent studies found that female scholars cite more female-authored work (Maliniak and Powers 2013; Mitchell, Lange, and Brus 2013).

The Smith et al. study (2020) hypothesized the following:

1. The proportion of assigned readings with female first authors is significantly lower than the proportion of female-authored publications in top journals.

2. Instructors from underrepresented groups assign more female-authored readings than instructors from dominant groups.

3. Older female instructors assign more female-authored readings than younger instructors, (older female instructors may have larger gender-homophilous networks, including the mentoring of junior women, making them more aware of women’s publications).

4. Instructors socialized in more gender-equal environments may assign more female-authored readings due to greater awareness of diversity and perhaps lower gender bias.

Their findings on the stated hypotheses show in hypothesis 1 that the proportion of assigned readings with female-first or only (“solo”) authors (18.5%) is lower than the female-authored proportion of top journal publications. Consistent with hypothesis 2, female instructors assign more female-authored readings (holding all variables at their observed values) 15.4% of readings assigned by male instructors are predicted to have female-first or solo authors, contrasted with 24.4% of readings assigned by female or mixed-gender instructors. Moreover, although the proportion of readings with female-first authors rises as a function of publication year, the gap between male and female instructors is not closing over time. Race was statistically significant among male instructors, whereas gender disparities in assigning work by women were observed only among white instructors. Nonwhite male instructors assign female-authored work at rates that are indistinguishable from those of female instructors (both white and nonwhite). Men and women born in countries that are less gender equal, according to the United Nations (UN), assign fewer readings authored by women (UN Development Programme 2017).

One of the implications of underrepresentation in syllabi is that it can affect female graduate student retention rates (Hardt et al. 2019, Smith et al. 2020). Additionally, male, and female students who rarely see women’s research may become less likely to cite women, thereby developing or reinforcing gender biases related to the quality of women’s research. For future research, the generational cohort over time may reflect tends toward greater gender representation. This line of research adds to the breadth of influences affecting the use of citations and their counts.

**Journals, Editors, and Citation-Related Research**

Scholarly journals are the heart of this academic reward system, which gravitates around hierarchies reflected by citation-based indicators such as the Journal Impact Factor. In this system, top journals attract what are believed to be the most important papers, which in turn,
bestow symbolic capital upon authors. Bourdieu’s (2004) article examines repeat authorships within scientific journals—authors who publish repeatedly in the same journal, especially high-status journals—as a specific phenomenon that reflects the Matthew Effects in science. The Matthew Effect (Azoulay et al., 2014) is a term denoting processes by which privileged scientists accrue further advantages and rewards solely by virtue of their status.

Experience with the peer review system in each journal—whether as an author or peer reviewer—helps develop tacit knowledge to successfully navigate that system in the future. Since innovators tend to repeat or emulate successful outcomes, this underpins incentives to focus on exploiting successful niches, instead of exploring new terrain (March 1991; Audia & Goncalo 2007). Although there is little evidence to suggest that editors are more likely to accept manuscripts written by men compared with women (e.g., Borsuk et al. 2009; Blank 1991), there is compelling evidence to suggest that women on average generate higher quality work during the peer review process (Hengel 2017), in part because reviewers and editors appear to impose higher standards on women (Card et al. 2020). Ultimately, this process could push women to “spend too much time rewriting old papers and not enough time writing new papers” (Hengel 2017:3).

Peer review is a core value and method of quality control in psychological research, academic psychology, and other disciplines, but very little is known about the peer-reviewing behavior of ethnic minority reviewers. Blume and Schmarling (2016) conducted an examination of 76 journals that used the JBO system over a 10-year period. They used hierarchical linear models, the percentage of review invitations extended to ethnic minorities was found to significantly increase over time: Initially, an estimated 8.34% of review requests were made to ethnic minority reviewers, and that percentage increased an average of 0.41% per year.

In addition, ethnic minority reviewers were significantly less likely to refuse a review request than ethnic majority reviewers. Results are discussed in terms of perceived pressure to demonstrate scholastic influence and the disproportionate service burden often borne by ethnic minority psychologists. For many ethnic minority scholars, there is concern that their scholarship is often devalued by colleagues as not in the mainstream or of high-impact journals (Bernal & Villalpando 2002; Evans & Cokley 2008; Thompson 2008), and certainly there is evidence that ethnic minority scholars fair less well than ethnic majority scholars on student teaching evaluations (Reid 2010). They hypothesized that self-identified ethnic minority reviewers would be less likely to refuse a review request than those who do not self-identify as an ethnic minority. Also, they expected that increases in reviewer burden would be evident in significant increases in declines to requests by all reviewers (Reid 2010). On the other hand, the proportion of requests to self-identified ethnic minority reviewers was lower than the proportion of ethnic minority psychologists in academia (13% in 2009 to 2010 Pagano, Kohout, & Wicherski 2010).

The blind review process does not disclose either reviewer(s) or submitter(s). It is expected that all journals hold reviewers accountable for being knowledgeable enough about the field to be able to identify when a manuscript omits and misrepresents the scholarship in the field and/or makes empirical claims without data or analysis. There is also the expectation that everyone who
submits to a journal to be held to consistent standards during the submission, review, and publication process.

The Zong et al. study (2019, 2020) investigates whether open peer review can improve citation count. Open Peer Review (OPR) is an umbrella term for many different models (Elsevier 2018; Foxe and Bolam 2017). In most cases peer review reports (signed or anonymous), together with authors’ and editors’ responses, are published alongside the article (signed or anonymous). The types include when open identities (both reviewers and authors are known to each other), open peer reports, open participation, open interaction, open prereview manuscripts, open final version commenting and open platforms. OPR has potential advantages, which include high-quality reviews (Sconfienza and Sardanelli 2013), giving credits to reviewers (Edmunds 2013; Boldt 2011; Sadeghi et al. 2019; Falk Delgado et al. 2019), helping young scientists (who can learn about the reviewing process of a manuscript) to prepare high-quality works (Rampelotto 2014; Dougherty and Dreher 2007), and improving accountability of editors and reviewers (Prager 2019). However, evidence regarding these potential advantages of OPR is inconclusive (Resnik and Elmore 2016). Several previous empirical studies indicated that OPR can improve the review quality (Leek et al. 2011; Bruce et al. 2016; Walsh et al. 2000; Kowalczyk et al. 2013, 2015).

Articles published in *Peer Journal* during 2013 and 2015 were chosen as the data set. Manuscripts submitted to *Peer Journal* are reviewed under the traditional single-blind peer review. It is “open” in the sense that reviewers can sign reviews, and authors can optionally share the review history (including reviewers’ comments, authors’ responses, editors’ decisions, any earlier versions of the manuscripts, etc.) when the manuscript is published. When authors choose to make their review history public, then *Peer Journal* creates a dedicated page for their publication’s history with the editor in charge, reviewers’ comments, any earlier versions of the manuscripts and authors’ responses (Peer Journal 2019b). That is, *Peer Journal* is somewhat unique in that it is all optional—the author and the reviewer have a choice (Peer Journal 2019c).

Two categories of the articles were generated: articles with closed peer review history and articles with open peer review history. A propensity score matching with the radius matching method was performed using 14 confounding variables (Zong, 2020). Their results suggest that open peer review can improve citation count, and that the best practice for open peer review might be a compromise open peer review. A closed peer review can be a single-blind review, a double-blind review or even a triple-blind review. In all closed peer reviews, the author of a manuscript does not know who the reviewers are. The closed peer review, especially the single-blind review, is the most common form of peer review among scientific journals (Wiley-Press 2018). In addition, despite the potential advantages of gaining credits in OPR, reviewers might be reluctant to participate in OPR (Groves and Loder 2014; Lee et al. 2013; van Rooyen et al. 2010; Nobarany and Booth 2015) because the publicly critical comments are a risk (Liu 2007; Janowicz and Hitzler 2012).

The confounding variables were characteristics of the articles. Previous studies indicated that title length (Guo et al. 2018; Gnewuch and Wohlrabe 2017), abstract length (Hafeez et al. 2019), number of authors (Ahmed et al. 2016), number of affiliations (Azer and Azer 2016), keyword
count (Uddin and Khan 2016), number of references (Didegah et al. 2018), page count (Hafeez et al. 2019), days since publication (Clements 2017), preprint (Davis and Fromerth 2007; Metcalfe 2006) and funding (Patience et al. 2017), may have effects on the outcome variable (citation count).

The results showed that, compared to articles with closed review history, articles presented with open peer review could be expected to gain more citations. This finding was to some extent consistent with the previous study conducted by Bornmann et al. (2010), which found that OPR can enhance the quality of a manuscript by increasing reviewers’ and authors’ accountability (Vercellini et al. 2016; Edmunds 2013), thus improving the citation count of the manuscript. That is, open identities and open review reports encourage reviewers to be more thorough in their review process (Ross-Hellauer and Gorogh 2019).

Zong et al. (2019, 2020) showed that approximately half of the respondents read the pre-publication history (including review reports, authors’ responses, and editors’ comments) published alongside the articles of the two journals working with the OPR. Besides reviewers’ comments, authors’ responses and editors’ comments, the webpages of peer review history on the Peer Journal website provide any earlier versions of the articles. This allows readers to see the scrutiny process and provides the opportunity for readers to observe what revisions were made (knowing what questions have been raised and how the authors have addressed them) based on the reviews (Baggs et al. 2008; Taylor 2018).

On the other hand, journals should provide a choice for reviewers to sign (or not sign) their names on the review reports and inform the reviewers that their review reports (signed or not signed) will be made public if the authors of the manuscripts they reviewed choose to publish the review history when the manuscripts are published. For researchers, it is a wise choice to publish review history on the condition that the journal they are working with offers optional OPR. Another compromise OPR is to mandate public peer review history, including reviewers’ comments (reviewers have a choice to sign or not), authors’ responses, editors’ declinations, and any earlier versions of the manuscripts, etc.

Several participants argued that OPR represented a form of “democratic practice” in science (Pharaon 2007) and ensured that reviewers are honest, more thoughtful, and less likely to exhibit ad hominem and ideological bias (Mulligan et al. 2013), thus strengthening scientific exchange (Pöschl 2010; Groves and Loder 2014) and quality reviews (Janowicz and Hitzler 2012). In contrast, others argued that OPR encouraged reviewers to be less critical (especially if the paper is written by a senior authority) (Mulligan et al. 2013; Khan 2010), thus making the peer review worse (Ross-Hellauer et al. 2017). Zong (2020) has noted that they are working with a number of journal editors to collect submission data and analyze these for their gender and methodological content. We invite other journal editors to participate in this project; the results can form the empirical basis for a broader discussion within the discipline. We highlight this study and open peer review as one of a variety of options that can have positive outcomes for research quality and reduce the bias experienced by some researchers.
A Litany of Other Contributing Factors

The Matthew Effect (Azoulay et al. 2014) is a term denoting processes by which privileged scientists accrue further advantages and rewards solely by virtue of their status. These processes are at odds with Merton’s (1973 [1942]) norm of universalism—the notion that scientists and their work should be judged and rewarded irrespective of their personal or social characteristics—as well as contemporary beliefs about meritocracy and fairness. Yet, research scholars are more likely to invoke particularistic characteristics of authors (e.g., gender, institutional status) as decisive information under conditions of uncertainty (Long & Fox 1995; Pfeffer et al. 1977), such as at the frontier of new scientific research (Cole 1983). This pattern is particularly the case in evaluative settings, as researchers are often influenced by the social status of an author.

For example, when researchers receive high-profile awards, their previous publications receive a boost in citations (Azoulay, Stuart & Wang 2014), which also causes intellectually proximate scholars to be crowded out of the research area (Reschke, Azoulay & Stuart 2017). Prestige-garnering publications in high-profile journals can function like similar public adornments of status on researchers. More established scholars also tend to enjoy professional advantages with social and intellectual networks, further helping them bear and expose their work. In turn, research tends to reproduce itself in both ideas and personnel (Bourdieu 1988). The tacit knowledge, social networks, experience, and reputations that researchers develop over time entail considerable professional and intellectual advantages. Merton (1973 [1942]) labeled science as a gerontocracy. Finally, success is also conducive to increased specialization in the future, as people tend to exploit and expand upon successful established niches in science, as opposed to exploring new terrain (March 1991; Audia & Goncalo, 2007) or establishing new subfields (Heinze et al., 2013).

However, the intellectual and professional devaluation of subfields and topics are overrepresented by women and people of color (Ferber & Nelson, 2017). Social networks, hierarchies in tenure, promotion and institutional status, underrepresentation of people of color and women graduate students in elite departments are other factors (Weeden, Thébaud & Gelbgiser, 2017), as well as conscious or unconscious biases against the evaluation and engagement of people of color and women serving as possible explanations for disparate rates of citations. Examples of selection biases result when an author’s ability to be published multiple times in the same journal is seen as indication of talent. Within the individual careers of scientists, citation impact steadily declines with each repeat authorship. This suggests that authors may be well served by diversifying their publishing portfolio.

A litany of other factors affecting citation counts include the following:

- Paper page count, (on expectation that longer papers contain more material and are thus more likely to contain material that others find worth citing (see Østby et al. 2013);
• The number of references that the paper made to other sources (based on expectation that larger number of references reflects the degree to which a paper is related to other research and thus reflects the potential of the paper to be cited (see Wang et al. 2011);

• Theoretical paradigm: realist, liberal, Marxist, constructivist, nonparadigmatic, or atheoretical;

• Age of publication (years since publication);

• A square term to account for the possibility that the effect of age changes with time;

• Author tenured:

  • Tenured female;

  • At an R1 institution or not;

• Co-authored: usually coded if the publication has more than one author;

• Epistemology;

• Noted as positivist: theoretical or empirical propositions are testable, etc.;

• Ideational: use of ideational factors such as where ideas, beliefs, perceptions, learning, norms, identity, knowledge, or personality traits play a central role;

• Material: if “material variables,” which are non-ideational and refer to ascriptive characteristics of actors or the structures in which actors are embedded;

• Issue area: international security, international political economy, human rights, environment, health, IR theory, US foreign policy, comparative foreign policy, history of the IR discipline, philosophy of science, international law, other, general or nonspecific, international organization, methodology, comparative politics, American politics, or political theory;

• Methodology: whether the study uses quantitative statistics, qualitative case studies, formal modeling calculus, game theory, spatial modeling, or some other methodological approach; and

• Journal of publication: Many articles can employ one or more of the following types of methods: quantitative, qualitative, formal modeling, counterfactual, analytics nonformal conceptual, descriptive, policy analysis, and experimental.

From our review of the extant literature, analyses examining the biases toward some scholars results in a multiplicity of factors and interacting factors that cross professional experiences and
status, race, ethnicity, gender and sexual orientation, fields of study and relation of one’s areas of research relative to the mainstream body of research. Our last section will focus on works that take their understanding of citation-related research and provide alternatives to minimize sources of biases.

**Recommended Changes to Address Citation Biases**

In examining the extant research on citation-related research, professional development, and its effect on various disciplines, we have found that there have been efforts to address known biases and broaden the flow of research to the broadest segments of the research community. For example, the field of economics is showing cognizance of these gender issues in the profession and field, with several recent works and opinion pieces identifying problems and solutions to gender equity issues in the discipline (e.g., Hengel 2017; *The Economist* 2017; Wolfers 2018; Sarsons 2018). This trend of increasing concentration of influence in leading journals runs counter to most other fields in contemporary science, which are instead seeing trends of decreased concentration of citations in top journals (Lozano, Larivière & Gingras 2012).

Political scientists have been asked to reflect on their own citation practices to ensure that their references are consistent with gendered distribution of research in their area. Likewise, journal editors can ask peer reviewers to explicitly consider whether article bibliographies are representative, including the distribution of author genders. Some journals have gone farther, explicitly evaluating the gender balance of article bibliographies and encouraging authors to remedy gendered citation gaps by providing additional space to do so (*International Studies Review* 2018). APSA sections that sponsor journals need to evaluate whether the publications provide ample descriptive representation of section members. In addition, those that select journal editorial teams should pay attention not only to their diversity, but also to their plans for addressing potential citation biases. Fortunately, tools, including the Gender Balance Assessment Tool (Sumner 2018), can help political scientists quickly and easily evaluate gender balance in their bibliographies. Finally, while women researchers have received the majority of attention with regard to examining citation biases, other researchers such as persons of color and LGBTQ scholars warrant attention and more detailed examination as to the extent and nature of biases.

More general suggestions directed to all faculty can include informing all of their students, especially women and persons of color of the value of self-citation. If self-citation is a common and conventional practice, and that we know it is, then these students need to be encouraged to advocate for themselves and their work. In addition, faculty should also make students aware of the benefits of co-authorship across gender lines since collaboration may be one way to increase the visibility of one’s scholarly work. At the same time, the selection of one’s co-authors should be based on research abilities and collaborative qualities. Nevertheless, the real advantages of co-authoring should be known. Finally, faculty can emphasize the value of networking by helping their students make connections, especially with notable members of the field. This is especially important for women in a field that is disproportionately dominated by men. The most important recommendation for departments is to take the existence of gender bias seriously when evaluating female scholars for promotion and review. A large portion of the gender citation gap,
Therefore, could be narrowed if men and women were made aware of these patterns and encouraged to be more gender neutral in who they choose to cite. One suggestion for editors is simply to monitor the ratio of male to female citations in the articles they publish.

Our attempt to lay out the extant research on citation-related research and professional development is quite extensive. The underlying question is how to evaluate and determine the contribution/influence of research on the scholarly production of knowledge. Our examination of this area clearly highlights the scope of determining and assessing how to calculate citations (i.e., sources, methods to do so, understanding possible biases, and contributing factors for potential unfairness, and uses) require more sophisticated thinking and complex methodological techniques beyond simple citation counts. Implicit biases about a researcher background, institution of training, primary areas or subfields of one’s research interests, epistemological foundations, publication outlets and protocols, and departure from mainstream conventions are very much integrated into citations and their usage. Usage of citation is increasing as numbers of services are available and readily user friendly. This pattern is likely to continue, and some researchers are being more strategic in creating Google Scholar pages to enhance their visibility and counts. Departments and institutions place high value in making decisions about the recruitment and retention of researchers. Thus, greater understanding as to the bases of algorithms to construct citation counts, appropriate uses and interpretation, and areas of systematic biases become more the responsibility of units, institutions, and other decision-makers to understand fully the nature of citations and appropriate applications.

Journals and Editors and Citation Practices and Guidelines

One of our working group’s objectives was to explore how political science journals and their editors deal with the citation aspect of doing research, assigning reviewers, and evaluating manuscript submissions. Our working group members were tasked with summarizing some of the practices that journals and editors are using to address historical citation inequities in order to develop some shared guidelines and recommendations for current and future editors to consider implementing. We did this by talking to several editors, past and present, to get a view from the inside, and by taking a broad look at the submission guidelines from nearly 70 political science journals. Overall, we find that some of the strategies developed in the IR journals have yet to diffuse across other subfields. Further, through our conversations, we believe that there are some actionable practices that most journal editors can implement easily to address citation inequities. Many of these actions will improve the quality of published work and will make the editing process easier, not more difficult, for the scholars who manage our journals.

Editorial Experiences with Addressing Citation Disparities

We spoke individually by phone with 10 former and current journal editors to get input from the individuals who have direct experience with the review process. We promised anonymity and an informal conversation in order to elicit candid thoughts on the citation gap issue. These conversations produced some new ideas and some feedback on existing policies. The editors had
a wide range of experience with both small subfield journals and general discipline journals. The editors we spoke with were evenly split between men and women, and there were about 50% scholars of color.

Most of the editors we spoke with did not enter the job with the “citation gap” in the front of their minds, and none of the editors we spoke with received any guidance or training on the issue when they took over. All the efforts to address these issues came from the editors themselves, at least those who were aware of it. One editor described long and detailed conversations with the incoming editorial team about different citation gaps in their subfield. These conversations led the team to reach out to another journal that had publicly posted policies about reducing the citation gap. But beyond these two cases of formal, team-wide prioritization, most of the efforts we heard about were at the individual-editor level. The degree of awareness among the persons interviewed seems to be related to their own personal experiences and professional network communications. This aspect of citation disparities is one in which the teams of journal editors could explore and discuss now and into the future. Thus, there does appear to be a need for a systematic set of guidelines or best practices around this issue for incoming journal editors. Indeed, nearly everyone we spoke with recommended it. One area of systematic examination of journal practices, especially for a new team of editors, is the need to include the extent and nature of this aspect of the profession in orientation materials and discussions.

Individual attention to address the issue was varied and idiosyncratic. One editor who did have the citation gap issue in the front of their mind described the poorer quality of work they received due to submissions not citing women and scholars of color. Because the editor was familiar with the oldest and newest work in this subfield, and was mostly assigned manuscripts within their subfield, they could see clearly how the citation gap affected theoretical and empirical quality. They addressed this in a number of ways, including during the “decline” or “revise and resubmit” stages of the process. They also spoke about finding reviewers who would be able to address the missing work in areas rich with work by women and scholars of color. A second editor also spoke about using this last strategy, purposefully assigning reviewers who would identify and address any citation gaps in submitted manuscripts. They, too, viewed manuscripts that ignored work by women and scholars of color to be deficient. For them it was an issue of research quality. Other editors we spoke with, while sympathetic to the issue, did not have any specific strategies for dealing with this issue. They had heard of other journals’ efforts to “nudge” authors into citation checks or of the existence of the scholarly databases like Women Also Know Stuff, but they had not incorporated anything systematic into their workflow. They did seem interested in learning more about resources that would help reduce the citation gap and make their job easier.

A few initial takeaways emerge from these experiences:

- First, one editor praised the growing use of field editors and editorial teams as one way to help mitigate the citation gap issue. Specialization allows editors to find reviewer pools that match each paper better. Being assigned a limited number of manuscripts from their specialized field allows editors to move away from the traditional use of single-editor-
networks, the authors’ bibliography, and journal-specific reviewer databases when they are overseeing a manuscript from an unfamiliar area. All three of these traditional editorial practices can reproduce citation inequities.

- Second, it is clear that the editors we spoke with were not given much guidance or training about this issue, though all appeared to welcome such guidance. All spoke about the overwhelming number of issues they need to confront as new editors and were open to guidelines on addressing citation gaps.

- Third, among the editors who actively engaged with citation practices in their workflow, there was a strong belief that addressing these gaps in manuscripts improved manuscript quality significantly. This is important because another editor noted their only reservation about being more proactive on this issue was its potential effect on research integrity. They worried that authors would simply “game” the citations to satisfy journal guidelines. This was not the experience for editors who engaged in this work. They were adamant that their attention to citation gaps improved the final manuscripts significantly, not only in terms of additions to the citations but also in where and how they were deployed. Again, policy changes were more the result of initiatives by particular editors than the result of any systematic practices.

- Finally, some of the resources that proactive editors used actually made their overall job easier not more difficult. That is, in looking for ways to alleviate the lack of citations to scholarship by women and scholars of color in many manuscripts, editors encountered new resources for finding excellent reviewers outside of their traditional networks. We discuss this in more detail next.

The editors all expressed the need for wider and deeper reviewer pools. Wider in the sense of incorporating more individuals from historically excluded groups, and deeper in the sense of reaching out to more individuals from all groups in the discipline. They universally agreed that increasing reviewer diversity would contribute significantly to citation diversity. While some of the editors expressed concerns that addressing citation gaps could potentially make their jobs and the jobs of some reviewers more difficult, others explained that the opposite was true. A number of editors specifically brought up some resources that made their jobs easier. The Women Also Know Stuff (WAKS) and the People of Color Also Know Stuff (POCAKS) searchable expert databases were specifically highlighted as valuable resources for widening and deepening reviewer pools. Editors can search by research topic and find women and scholars of color who may lie outside of traditional reviewer networks. One editor noted that WAKS was the first place they went to when beginning to build a reviewer pool for papers. They noted that finding a single lead on that website led them to co-authors of that scholar, people that scholar cited, as well as scholars who cited their work. Within a couple of minutes, an entire set of outside-network reviewers can be put together. The same is true for the POCAKS database and the wide use of organizational caucus websites to look for subfield experts. We discuss the transformative potential for these kinds of databases in our recommendations but using these types of resources can help mitigate both the specific citation gap issue and the broader issue of finding reviewers.
These workflow efficiencies will make broader implementation of the guidelines much more likely.

Many of the editors discussed proactive best practices that could help diminish various influences on the citation gap. Some discussed the linkages between submission diversity and citation diversity. They argued for journals encouraging more manuscript submissions from women and scholars of color as a way of shrinking the gap. This could take the form of general calls, from the journal, for more research from areas studied by women and historically excluded scholars (Garcia and Hancock Alfaro 2021). One editor discussed the rapid success of bringing in a field editor from a neglected area to help encourage more submission from women and scholars of color in that subfield. Another pointed to the importance of specific language in calls for special issues that encourages historically excluded scholars, and teams involving graduate students and junior scholars, to submit proposals. One editor noted the importance of getting these calls out to specific networks of scholars whose work has been ignored and marginalized by many of these journals. The editors also noted that these efforts are likely to be undertaken by people who are already attending to the citation gaps in their weekly workflow, but that these are some of the proactive ways in which they achieved results. Again, some of these efforts make the jobs of editors easier, particularly the idea of bringing in field editor who can manage increased calls for research diversity in submissions. The editors that reached out to a more diversified group of field editors indicated that it was trusting these individuals to do their work that changed submission patterns without adding to their own work. In this case it was mostly about delegation instead of getting the entire editorial team on board to new ways of doing things.

Finally, the use of citation checks in the review process received a good deal of attention during these conversations. Citation checks allow authors to submit their bibliography to an online system that returns an estimated gender or racial/ethnic name check of cited authors (Fulvio 2020; Sumner 2018). As we discuss below, most political science journals do not have a publicly available statement about the citation gaps on their submission pages. Those that do typically encourage authors to consider the literature on citation gaps and address them before submitting. In fact, one of the recurring messages we heard from editors was the value of reminding authors about the existence of citation gaps and encouraging them to check their citations before submitting. Reminders like these can change individual practices at the margins which can have aggregate effects. Editors were consistently in favor of “nudges” like these to encourage authors to be more mindful and proactive about their citation practices within the context of the research on these gaps. This can be done formally at the submission guidelines stage, the revise and resubmit stage, and the rejection stage of the review process. We discuss the prevalence of these formal practices in the next section.

Review of Submission Guidelines

With the help of a paid graduate assistant, Erika Vallejo at Michigan State University, we scanned the submission guidelines of 68 political science journals to see if/how journals were representing the citation gap research to submitting authors. The list of journals to include came from suggestions by the members of the committee as well as publicly available top-50 lists of
political science journals. It includes all the most widely read and cited general and subfield journals in the discipline. The online searches took place between May 10, 2021, and July 18, 2021. We looked at the submission guideline pages and “Instructions for Authors” guides for each journal and looked for the terms “gender,” “citation,” “gap,” and “reference,” reading any section that stood out. We then carefully read the section of the submission guidelines that discussed reference practices to make sure we did not miss any other reference to citation gaps. We summarize the findings below.

Out of 68 journal submission guidelines and editorial instructions posted online that we reviewed, only 12 (17%) of them make any reference at all to a citation gap. Of those that do, all mention the “gender gap,” but just one specifically references a racial or ethnic gap. A couple of journals make oblique mentions beyond the gender gap but expresses this in very general terms like “underrepresented” or “disadvantaged.” For example, one journal notes “the well-documented biases that exist in academic citation practices, [and we] ask all authors to check their references to examine the extent to which they engage with researchers from underrepresented groups and diverse backgrounds.” Interestingly, eight of the 12 journals that mention a citation gap are international politics journals, which reflects an intensive effort in the IR field by individual leaders who pressed the issue for years to address this gap. We discuss these efforts below. The other four journals are a mix of subfield-specific journals. None of the top general journals that we reviewed had editorial instructions regarding the citation gap for authors to consider. Clearly there is much to be done in terms of journals at least mentioning the existence of a citation gap in the places where guidelines are most visible to prospective authors—journal submission pages. All journal statements are listed in the appendix for reference.

The International Relations field journals have taken the most public action, with many of them including reminders about the citation gaps in the literature and Asking authors to consider these citation gaps before submitting a manuscript. *International Studies Quarterly (ISQ)* has taken the most publicly proactive steps, including strong language in its submission guidelines (*International Interactions* has similarly strong language):

*We strongly recommend that authors check their references to ensure inclusion of authors from disadvantaged groups.* ISQ is committed to ensuring that scholars receive appropriate intellectual acknowledgement regardless of race, gender, class, professional standing, or other categorical attributes.

*Review of International Studies* includes the following language in its peer review policy, although it is within the larger section of instructions to authors:

*We also ask reviewers to consider whether adequate reference is made to the scholarly literature in the field. Recent studies have highlighted the possible under-representation of female and minority scholars in article citations. Review of International Studies is committed to ensuring that scholars receive appropriate intellectual acknowledgement through citations, regardless of race, gender, class, professional standing, or other*
categorical attributes. To that end, we ask referees to pay particular attention to the representativeness of citational practices manifested in all article submissions.

Guidance from other journals falls into the “take under consideration” category. *International Organization* advises, “As you may know, recent research has shown that there is a gender citation gap. We ask you to be mindful of this when preparing your submission.” *Political Analysis* advises, “Please consider the gender balance of your citation list and when considering contributing to your topic.” *Political Communication* takes this one step further, suggesting some resources for authors to use to widen their knowledge network:

The journal encourages submitting authors (and reviewers) to consider the composition of the cited authors or suggested works with respect to gender and minority representation. The journal encourages the use of e.g., the (searchable by topic/expert) WomenAlsoKnowStuff and PeopleOfColorAlsoKnowStuff data bases.

We subsequently discovered that the editors at *ISQ* have taken the formal step of reminding authors again at the R&R stage to consider the citation gap in their revisions and providing a target percentage as a goal.

*International Studies Quarterly* is committed to ensuring that scholars receive appropriate intellectual acknowledgement regardless of race, gender, class, professional standing, or other categorical attributes. Please pay particular attention to this issue when revising your citations for overlooked authors and literatures. You can easily check the gender-balance of your references by using the GBAT tool found here: [https://jlsurner.shinyapps.io/syllabustool/](https://jlsurner.shinyapps.io/syllabustool/). Given the persistent gender citation gap in international relations and the increasing number of female students and faculty in the discipline, we aim for approximately 30% female citations to ensure appropriate scholarly recognition.

This does not appear to be a requirement, but it does send a strong signal about how important this issue is to the editors. We are not aware of other journal editors who use this language but it’s possible that some do.

The diffusion of citation gap language in the IR journals’ submission guidelines is not coincidental. Much of the research in this area has been done by IR and Methods scholars (Dion and Michell 2020; Dion, Mitchell, and Sumner 2020; Dion, Sumner, and Mitchell 2018; Maliniak, Powers, Walter 2013) and there has been a strong push by individual scholars in the IR field to change the editorial practices of their journals in response to research citation disparities. This did not happen by accident or overnight. As we learned during our conversations, it started with IR scholars asking the journals to collect the data, studying the data systematically, and then implementing changes in response to what the data were saying via IR professional committees. Thus, this process took years of time and effort and has clearly been successful in terms of alerting IR scholars to the citation gap issues and changing some editorial practices.
To conclude, most journals do not appear to reference citation gaps and disparities at all in their submission guidelines. Those that do tend to come from the IR field, and the language they use varies from fairly neutral to proactively strong. Additionally, we do see variation in the strength of the language, the general focus on gender versus racial, ethnic, and geographic citation gaps, and the degree to which journals offer resources to authors. We cannot, however, assess the utility of this submission guideline language in helping to mitigate the citation gaps at specific journals.

**Recommendations and Suggested Guidelines**

Our review of the literature, publicly available journal policies, and interviews with editors suggest a number of concrete steps that editors can take to ensure broader citation practices that will create higher quality research.

1. **Develop journal statements and guidance about citation gaps**

   Journals are strongly encouraged to take the simple step of being upfront about the citation gaps in the discipline and asking authors to consider these gaps as they submit manuscripts for review. *These statements should be common across all of the journal submission pages.* We provide several example statements in the appendix, but our recommendation is to incorporate language that goes beyond the gender citation gap and includes historically excluded scholars, e.g., racial, and ethnic minority scholars, LGBTQ+, and Indigenous scholars (Garcia and Hancock Alfaro 2021), and language that encourages authors to check their references using one of the widely available tools (below). We would also recommend editors incorporate language into their standard correspondence to authors to remind them about these gaps and how to check for bias using online tools:

   a. [https://postlab.psych.wisc.edu/gcbialyzer/](https://postlab.psych.wisc.edu/gcbialyzer/)
   
   b. [https://jlsumner.shinyapps.io/syllabustool/](https://jlsumner.shinyapps.io/syllabustool/)

2. **Develop new tools to help editors broaden their reviewer pools**

   All of the editors we spoke to tie the citation gap to reviewers and reviewer pools in one way or another. Expanding the reviewer pool beyond people who are cited by authors, and beyond names in the editors’ network or the journal’s database is both helpful to mitigating citation gaps and also to making editors’ jobs better. *We recommend that APSA create a database of reviewers that editors can access, by research area and with demographic information that APSA already collects.* This would allow editors to look beyond their own networks and the networks of authors for reviewers. It would also help editors find new and willing reviewers, one of the most difficult tasks in the workflow of editors and their staff. This database would supplement other valuable resources that have been created to expand the expert pool. All these resources should be provided to incoming teams of editors so that they are encouraged to use them when finding reviewers. Every editor we talked to welcome this idea and encouraged APSA to undertake this
fairly straightforward step to helping journal editors find the best reviewers for submitted manuscripts. In the meantime, editors should be encouraged to use publicly available resources like those provided by People of Color Also Know Stuff and Women Also Know Stuff to expand their reviewer pools.

a. https://sites.google.com/view/pocexperts/home

b. https://womenalsoknowstuff.com/

One more straightforward thing journals can do is to strongly enforce the norm of asking reviewers who decline an invitation to review instead to suggest additional names of experts to invite as reviewers. Having reviewers help editors expand the reviewer pool will solve some of the same issues discussed above and should become a norm in the discipline that is framed as a simple and effective way to help mitigate citation gaps.

3. Use field editors and editor teams

We recommend the continued trend of bringing in field editors or using the editorial teams models for journals to guarantee that editors are reading and making decisions about manuscripts in their area of expertise. The editors we spoke with who were able to directly address citation gaps were the ones who were reviewing manuscripts within their specific subfield. This becomes more difficult to do at the general journals, where editors are tasked with reviewing many manuscripts outside of their field of knowledge, thus falling back on some of the editorial methods that can reproduce citation gaps. Using field editors or having larger teams of editors at the general journals will help alleviate this problem.

One other issue that should be explored by incoming editorial teams is the composition and training of the editorial assistant teams. As one editor noted, editorial assistants, particularly at the larger general journals, are often offering the first set of reviewer options for editors. This common reality requires that these assistants are brought into the workflow conversations about measures to reducing citation gaps in submitted papers.

4. APSA citation gap study for historically excluded scholars

We recommend that APSA undertake an empirical study of citation gaps for historically excluded scholars. Many of the changes we document here were brought about because (mostly IR) journals were compelled to collect and study the citation data in their journals. Studying racial/ethnic and other historically marginalized groups will be more difficult on a variety of dimensions but should be done. A funded empirical study of the citation issue similar to those done for gender gaps would help highlight other ways in which authors, by not citing researchers from underrepresented racial and ethnic groups, are erasing research done by individuals from groups that have been historically ignored. The continued dominance of American-centric (and European-centric) scholarship, even in an increasingly transnational world, is a product of the self-enforcing citation behaviors that tend to reproduce themselves. The absence of Indigenous,
Native, and people of color in citations, even in studies about their own communities, societies, and politics, reinforces this asymmetry of global knowledge production and consumption at the expense of excluding scholars who are from the Global South, traditionally excluded communities, and Indigenous people. Taken together these issues assume that racial and ethnic hierarchy are embedded in who is capable of not only producing knowledge, but also whose knowledge is worth citing and expanding. Various editors noted that these groups have largely been left out of the citation gap conversation because of a dearth of empirical work in this area. We believe APSA is the appropriate organization to help support this type of empirical research. In summer 2021, we conducted a study to examine demographic and citation information of more than 4,000 APSA members. We discuss that study and our findings in the next section.

Race/Ethnicity and Citation Counts of APSA Members

In the summer of 2021, we appended demographic information and citation information to a database of 4,077 APSA members. The dataset was provided by Natalie Masouka and her colleagues (2007) and augmented by the APSA research unit. The dataset contained information about each member’s title, the university at which they are currently employed, the institution at which they received their PhD, the year in which they graduated, their first and second fields of study, their gender, and their race/ethnicity. The compiled dataset also includes an indicator of whether the member has a Google Scholar page, their citation counts via Google Scholar if available, and citation counts via Scopus (Elsevier’s citation database).

Prior work has examined citation counts among members of the discipline primarily for the purpose of identifying the most cited scholars in the profession and highlighting departments where average citation counts of scholars are notably high (Klingemann 1986; Klingemann, Grofman, and Campagna 1989; Masuoka, Grofman, and Feld 2007; Miller, Tien, and Peebler 1996). Here, we employ citation data paired with APSA member information to conduct a preliminary analysis of differences in citation counts by the race and ethnicity of different scholars. While a growing body of research has considered citation gaps across academic disciplines with respect to gender (e.g., Dion, Sumner, and Mitchell 2018), our analysis here makes an important and novel contribution by examining citation gaps by race and ethnicity.

Characteristics of the Database

For some members, their race or ethnicity was provided by APSA. For others, the race or ethnicity was identified by the researchers. In total, 4.2% of members identify as (or were identified as) Asian, 2.9% as Black, 2.6% as Latino, 58.0% as white, 1.3% as multi-racial, and 20.1% as “other.” Race or ethnicity is not available for 10.9% of the members.

In total, members in the database came from over 200 universities. We identified these universities as either R1 (including high-ranking international universities, like Cambridge University, that do not get the formal R1 designation) or not. A total of 3,537 (86.8%) members report being at a R1 university, 540 (13.2%) reported being at another type of institution, and the remaining values were missing.
Members in the data represent a diverse array of discipline subfield specialties. As reported in Table 1, a sizable plurality (29%) reported American Politics as their first field. Approximately 25% identified comparative politics, and 22% identified IR. About 9% reported political theory as their first field, 5% reported public policy, 4% reported methodology, and no more than 2% composed the remaining fields of public administration, public law, and race, ethnicity, and politics.

Table 1. APSA Members by Discipline Subfield (Self-Identified First Field of Study)

<table>
<thead>
<tr>
<th>Field</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Politics</td>
<td>1,186</td>
<td>29</td>
</tr>
<tr>
<td>Comparative Politics</td>
<td>1024</td>
<td>25</td>
</tr>
<tr>
<td>International Relations</td>
<td>903</td>
<td>22</td>
</tr>
<tr>
<td>Methodology</td>
<td>153</td>
<td>4</td>
</tr>
<tr>
<td>Political Psychology</td>
<td>36</td>
<td>1</td>
</tr>
<tr>
<td>Political Theory</td>
<td>399</td>
<td>10</td>
</tr>
<tr>
<td>Public Administration</td>
<td>61</td>
<td>1</td>
</tr>
<tr>
<td>Public Law</td>
<td>64</td>
<td>2</td>
</tr>
<tr>
<td>Public Policy</td>
<td>220</td>
<td>5</td>
</tr>
<tr>
<td>Race and Ethnicity</td>
<td>31</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>4,077</td>
<td>100</td>
</tr>
</tbody>
</table>

Source of Citation Counts

We included in the dataset, when possible, the number of citations for each scholar as recorded by Google Scholar. Unlike Scopus, Google Scholar does not generate citation counts for all researchers; instead, scholars must opt into having a Google Scholar page for Google to make their citation counts available. The different sources for citation counts are important, because relative to Google Scholar, Scopus seems to systematically undercount the number of times a scholar is cited by an order of magnitude.

We found notable differences in rates of having a Google Scholar page by racial groups. As we illustrate in Table 2, approximately 56% of Black scholars in the database had a Google Scholar, while only 25% of Latino scholars had a page. Close to 35% of Asian scholars had a Google Scholar page, and about 41% of white scholars did as well. We also note that only 39% of women in the data had a Google Scholar page, compared to approximately 45% of men. Interestingly, we found that scholars not at R1 institutions were more likely (47%) to have a
Google Scholar page than those at R1 institutions (42%). What is more, scholars in more recent cohorts were less likely to have a page than those in older cohorts.

**Table 2. APSA Members’ Presence of Google Scholar Page by Race and Ethnicity**

<table>
<thead>
<tr>
<th></th>
<th>Asian</th>
<th>Black</th>
<th>Latino</th>
<th>White</th>
<th>Multiracial</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>65.28%</td>
<td>43.51%</td>
<td>74.58%</td>
<td>58.89%</td>
<td>64.41%</td>
<td>48.91%</td>
</tr>
<tr>
<td>Yes</td>
<td>34.72%</td>
<td>56.49%</td>
<td>25.42%</td>
<td>41.11%</td>
<td>35.59%</td>
<td>51.09%</td>
</tr>
</tbody>
</table>

Because such a sizable proportion of scholars in the data do not have a Google Scholar page, we therefore conduct subsequent analyses using Scopus citations. While Scopus seems to undercount citations, we do note that Scopus and Google Scholar citations trend in the same direction (Pearson $r$ correlation coefficient between the two citation counts is 0.75). Thus, we have confidence that most of the patterns we uncover using Scopus citations would be similar if we were to use Google Scholar data. Nevertheless, we raise the possibility that the lack of a Google Scholar page may be introducing downstream bias into who is more likely to get cited (and the causal arrow may go the other way as well). Future work should further investigate the factors that may be associated with whether a scholar has a Google Scholar page.

**The Distribution of Citations by Subgroups**

To illustrate citation patterns by race, we compare the mean number of citations across different subgroups. Black scholars have a mean number of citations of about 245 and Latino scholars have a mean number of citations of approximately 115, compared to white scholars, who have a mean of 690 citations (mean difference between Blacks and whites and Latinos and whites of 445 and 309, respectively, significant at $p=.000$). Asian scholars have a mean number of citations of 187. We find that women of color have an average of approximately 408 citations, compared to white women, who have an average of about 705 citations (mean difference of 297.5 significant at $p=.000$).

Citation counts also differ by subfield. American Politics scholars have a mean number of citations of approximately 766—the highest on average number of citations across subfields. The mean number of citations among those whose primary field is IR is 663. Political theorists have a mean number of citations of 613, and those who study comparative politics have a mean of 605.

Next, we compare in Table 3 the mean number of citations by race and ethnicity among scholars at R1 universities. On average, white scholars are cited more than others (although the large standard error suggests that a small subset of scholars is elevating the mean). We also see in Table 3 that white scholars at R1 universities have an average of 857 citations, compared to Black scholars who have an average of 288, Latino scholars who have an average of 399, and Asian scholars who have an average of 614. We also note that the lower levels of citations
received by racial and ethnic minority scholars relative to whites persists across discipline subfields, and the differences are especially stark at R1 universities.

Table 3. Mean Number of APSA Member’s SCOPUS Citation Counts by Race and Ethnicity

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>287.6602</td>
<td>445.3333</td>
<td>103</td>
</tr>
<tr>
<td>Latino</td>
<td>399.1887</td>
<td>555.0755</td>
<td>106</td>
</tr>
<tr>
<td>Asian</td>
<td>614.4545</td>
<td>1260.516</td>
<td>165</td>
</tr>
<tr>
<td>White</td>
<td>857.2758</td>
<td>1649.878</td>
<td>2,255</td>
</tr>
<tr>
<td>“Non-White”</td>
<td>484.4187</td>
<td>978.0626</td>
<td>1,132</td>
</tr>
</tbody>
</table>

Note: Analysis limited to scholars at R1 universities.

We also considered whether there were notable racial and ethnic differences in citation counts by cohort. We find some evidence of greater equality in citation patterns among younger cohorts. For instance, scholars of color in more recent cohorts—those earning their PhD in 2010 or later—have a mean of 172 citations, compared to whites in this group who have a mean of 242 citations (the difference is not statistically significant at conventional levels). Among scholars who got their PhDs in the 1990s, however, racial, and ethnic differences are stark. Scholars of color from this generation have a mean number of citations of 690, compared to white scholars’ mean of 1029. These trends could reflect biases that occur over the lifetime of scholars’ careers, or they could signal a move toward less bias in citations patterns among newer generations of scholars.

The Factors Associated with More Citations

We estimate using Poisson regression the number of Scopus citations as a function of gender, race or ethnicity, employment at a R1 university, discipline subfield, and year of earning PhD. As we can see in Table 4, there are several significant relationships that emerge from the multivariate analysis. Women, all else equal, had fewer citations than men. Identifying as Asian or white was associated with a greater number of citations, while identifying as Black, Latino, or multiracial was associated with fewer (although the effect among Latinos was not statistically significant). Scholars at R1 universities are cited more, as are those who achieved their PhD less recently—an unsurprising result given that more time in the discipline means more time to accrue citations.
Table 4. The Factors Associated with Higher Numbers of Scopus Citations

<table>
<thead>
<tr>
<th>Category</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>-0.315***</td>
<td>(0.0534)</td>
</tr>
<tr>
<td>Asian</td>
<td>-0.128</td>
<td>(0.112)</td>
</tr>
<tr>
<td>Black</td>
<td>-0.956***</td>
<td>(0.138)</td>
</tr>
<tr>
<td>Other</td>
<td>-0.372***</td>
<td>(0.0607)</td>
</tr>
<tr>
<td>Latino</td>
<td>-0.517***</td>
<td>(0.141)</td>
</tr>
<tr>
<td>Multiracial</td>
<td>-0.300</td>
<td>(0.197)</td>
</tr>
<tr>
<td>R-1 University</td>
<td>0.883***</td>
<td>(0.0727)</td>
</tr>
<tr>
<td>American Politics</td>
<td>0.0705</td>
<td>(0.0769)</td>
</tr>
<tr>
<td>Comparative Politics</td>
<td>-0.0134</td>
<td>(0.0788)</td>
</tr>
<tr>
<td>International Relations</td>
<td>-0.00336</td>
<td>(0.0807)</td>
</tr>
<tr>
<td>Political Theory</td>
<td>-0.0937</td>
<td>(0.0988)</td>
</tr>
<tr>
<td>Year of Ph.D.</td>
<td>-0.0184***</td>
<td>(0.00184)</td>
</tr>
<tr>
<td>Constant</td>
<td>6.708***</td>
<td>(0.125)</td>
</tr>
</tbody>
</table>

Observations 3,653

Note: Excluded racial category is white. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Conclusion

Our analysis, one of the few that has examined citation gaps by race and ethnicity, uncovers significant gaps in citations across members of the political science profession. In keeping with longstanding patterns of racial and ethnic inequality in the discipline, white scholars, across subfields and institutions, are cited at significantly higher rates than scholars of color. There is some evidence that these gaps are less severe among younger generations of scholars, but it is
possible that they may diverge over the course of scholars’ careers. Our analysis, therefore, puts in stark relief the need for members of the profession to be more deliberate and expeditious in working to remedy citation gaps, which reproduce inequalities in opportunities that drive professional success.

Race and ethnicity, of course, are not the only factors associated with citation counts. We also find that type of university and cohort are significantly tied to citation counts as well—factors that when interacted with race and ethnicity, may also exacerbate inequality. What is more, given the data we have available, we are not able to examine the potential effects of other potentially compounding factors, including tenured status, parental status, and concentration in certain fields of study—all factors that when interacted with race or ethnicity, may further contribute to citation gaps. Our results therefore point to a need for scholars to investigate further the causes and consequences of citation gaps. Our work here, however, is an important first step in illuminating inequalities. It also provides an important caveat to universities that use citation counts when evaluating scholars for promotion. Longstanding biases in the discipline may mean that scholars of color are penalized relative to whites when citation counts are given significant weight in making these decisions. ²

² Excellent research assistance and data analysis for this memo was provided by Melanie Sayuri Dominguez and Stefanie Pousoulides.
References


Østby, Gudrun, Håvard Strand, Ragnhild Nordås, and Nils Petter Gleditsch. (2013). “Gender Gap or Gender Bias in Peace Research? Publication Patterns and Citation Rates for Journal of Peace Research,


Zigerell, L. J. 2015). “Is the Gender Citation Gap in International Relations Driven by Elite Papers?” Available at https://doi.org/10.1177%2F2053168015585192.


Appendix A: Meta-Analysis References Regarding Research, Citations and Professional Development


Available at [www.apsanet.org/RESOURCES/Data-on-the-Profession/Dashboard/Membership/Organized-Sections](http://www.apsanet.org/RESOURCES/Data-on-the-Profession/Dashboard/Membership/Organized-Sections).


Clarivate Analytics: (2019) “Web of Science”.


Gingras, Yves, Vincent Larivière, Benoît Macaluso, Jean-Pierre Robitaille. (2008). “The Effects of Aging on Researchers’ Publication and Citation Patterns.” *PLOS ONE*, 3(12), e4048.


Grimson, J. (2014). “Measuring research impact: Not everything that can be counted counts, and not everything that counts can be counted.” In W. Blockmans, L. Engwall, & D. Weaire (Eds.), *Bibliometrics use and abuse in the review of research performance* (Vol. 87, pp. 29–41). Wenner–Gren international.


Harzing, Anne-Wil 2010 *The Publish or Perish Book: Your Guide to Effective and Responsible Citation Analysis*+ Melbourne, Australia: Tarma Software Research.


Howard, Robert M. (2018.) “SPSA Membership and Gender.” Personal correspondence, October 12.


suggested reviewers in journals operating on open or closed peer review models. Paper presented at the 7th International Congress on Peer Review and Biomedical Publication 2013, Chicago, USA, 8–10 September.


Liu, S. V. (2007). “Why are people reluctant to join in open review?” *Nature*, 447(7148), 1052. [https://doi.org/10.1038/4471052d](https://doi.org/10.1038/4471052d).


Mershon, Carol, and Denise Walsh. "How political science can be more diverse." PS, Political Science & Politics 48, no. 3 (2015): 441.


Proceedings of the National Academy of Sciences of the United States of America, 101(suppl 1), 5200–5205.


Østby, Gudrun, Håvard Strand, Ragnhild Nordås, and Nils Petter Gleditsch. (2013). “Gender Gap or Gender Bias in Peace Research? Publication Patterns and Citation Rates for Journal of Peace Research,


Sdellavi. “2018). “Fact 3 on Top Journals: Given facts 1 and 2, acceptance rates in top journals keep declining, are at 2.5% for QJE, 3% for RES, 4% at JPE, 5-6% at AER and EMA” Apr. 25 [Tweet].


Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Stephen Curry, S. H., Jones, R., Kain, R., Kerridge, S.,


Zigerell, L. J. 2015). “Is the Gender Citation Gap in International Relations Driven by Elite Papers?” Available at https://doi.org/10.1177%2F2053168015585192.


