DRAFT

Men set their own cites high:
Gender and self-citation across fields and over time

Authors:
Molly M. King* - Department of Sociology, Stanford University, Stanford, CA
Shelley J. Correll - Department of Sociology, Stanford University, Stanford, CA
Jennifer Jacquet - Department of Environmental Studies, New York University, New York, NY
Carl T. Bergstrom - Department of Biology, University of Washington, Seattle, WA
Jevin D. West - School of Information, University of Washington, Seattle, WA

ABSTRACT

Self-citation is a common practice in scholarly publication; we find that nearly 10% of references are self-citations by a paper's authors. However, in almost all academic fields, men cite their own research papers at a higher rate than women do. Using a dataset of 1.6 million papers in the scholarly database JSTOR, we present results across many academic fields. Despite increased representation of women in academia, the gender gap in self-citation rates has widened over the last 50 years. These findings have important implications for scholarly visibility and potential consequences for academic careers.

Acknowledgments and credits:
This work was supported in part by a NSF Graduate Research Fellowship (grant DGE-1147470) to MMK, a John Templeton Foundation Metaknowledge Network grant to JDW and CTB, and a generous gift from JSTOR. The authors thank the Mathematica Stack Exchange community for assistance in creating Figure 2.

Grant numbers and/or funding information:
This work was supported in part by a NSF Graduate Research Fellowship (grant DGE-1147470) to MMK, and a John Templeton Foundation Metaknowledge Network grant to JDW and CTB.

***Corresponding author (name, address, phone/fax, e-mail)**
Molly M. King
Stanford University
Department of Sociology
450 Serra Mall
Building 120, Room 160
Stanford, CA 94305-2047

Email: kingmo@stanford.edu

INTRODUCTION

Scholarly publications are among the most important indicators of academic achievement. While the quantity of papers authored certainly matters, simple publication count is not the only important metric. The reputation of the journal in which a paper is published (often gauged using impact factors), along with the number of citations that a paper receives (i.e., other articles that reference that particular work), are together often seen as proxies for a publication's importance and influence.

Self-citation may have a consequential impact on scholarly careers by both directly and indirectly increasing an author's citation counts. Each additional self-citation yields an additional three citations (though not necessarily to the same paper) from other scholars over a five-year period (Fowler and Aksnes 2007). Given the importance of metrics of scholarly influence in academic hiring, tenure and salary decisions, examining gender differences in citation patterns may shed light on persisting gender discrepancies in faculty hiring and promotion. More broadly, academic publishing provides an illustrative case for gender differences in evaluation metrics and workplace advancement.

Papers authored by women receive fewer citations than do papers by men, even controlling for tenure status, institution, and journal (Larivière et al. 2013). Fewer citations to female-authored papers could be due in part to gender differences in self-citations (when an author cites his or her own previously published work). Research analyzing 12 journals in the field of international relations from 1986-2000 showed men cite their own papers more than one and a half times as often as women (Maliniak, Powers, and Walter 2013).[1]

To date, studies of self-citation have been few in number and confined to a limited number of disciplines and a relatively small number of papers. Here we examine gender differences in self-citations across 24 broad academic fields with hundreds of subfields and several million scholarly papers, with over a million self-citations. We further examine how the gender ratio self-citation patterns changed over time.

METHODS

The JSTOR corpus

JSTOR is a not-for-profit digital collection of scholarly documents ranging in time from the mid-sixteenth century to the present day. The JSTOR collection includes over eight million individual documents and over four million research articles, of which 1.8 million are linked by citation to other articles in the collection. Nearly 1.7 million of those articles make up a single giant component – meaning any document in the set can be reached by following citations (forward or

[1] A study of five archaeology publications also found that men tend to cite themselves slightly more often than women. This trend was not statistically significant - leading the author to conclude there was no gender difference in self-citation, even though the lack of significance could instead be due to the small sample sizes (Hutson 2006).

backward) from any other document. This single component includes papers back to the seventeenth century. However, for this analysis, we included only papers written in or after 1950, reducing our analytical dataset to 1.6 million papers. We focus on these documents, which we call the "JSTOR network dataset", because they are amenable to citation network analysis.

Mapping the hierarchical structure of scholarly research

A prior analysis (West et al. 2013) used the hierarchical map equation (Rosvall and Bergstrom 2011) to create a nested hierarchy of all documents in the network dataset. This hierarchical classification revealed the structure of fields, subfields, and ever-finer partitions down to the level of individual research topics. The algorithm by West et al. (2013) determined the boundaries between groups at each level of the hierarchy. We manually assigned names to the field, subfield, and research topic groups that the algorithm revealed.

Determining gender of authors

To assign gender to first name, we used the methods of West et al (2013), which relied on US Social Security Administration records[2] to provide information about first names and corresponding gender. (We therefore followed the US Social Security data in acknowledging only two genders.) We assume that we can confidently assign gender to author if the author's first name has the same gender at least 95% of the time in the Social Security database. Authors with first names that are associated with both genders, such as 'Jody' or 'Shannon' were dropped from the analysis. We assigned gender to authors' names that appear in the top 1000 most popular names in any year from 1879 – 2012.

We extracted the first names of all authors in the JSTOR network dataset, discarding those who list only initials. The exclusion of authors with only first initials may exclude women authors disproportionately, particularly in early eras when women may have been more likely than men to publish with initials to avoid potential discrimination. Since in any given era, androgynous names are more likely to be women (Lieberson, Dumais, and Baumann 2000), this may slightly downwardly bias our assignments of women. Similarly, we were unable to classify names that were not in the top 1000 for any year from 1879 – 2012. As a result, authors of some nationalities may be underrepresented in our data set. In a few rare cases, national differences may cause misleading assignments for non-US authors (e.g. 'Andrea' is typically a woman's name in the US but a man's name in Italy).

An *instance of authorship* consists of a person and a paper for which the person is designated as a sole or co-author. There are 3.6 million authorships in the JSTOR network dataset; of these, we were able to extract a full first name for 2.8 million authorships (77%) associated with 1.5 million papers. Of these 2.8 million authorships with full first names, we were able to confidently assign gender to 73.3%. The remaining authorships involve names not in the US social security lists (24.3%), or names associated with both genders (2.4%). The final data analyzed include all papers where we know the gender of one or more authors.

[2] Available at http://www.ssa.gov/oact/babynames/

Self-citations: an author-to-author approach

The inability to disambiguate authors is a challenge in this analysis. Disambiguation would highlight ties between papers by identifying when the same name belongs to the same individual across different authorship instances. We could have fully disambiguated the authors on a small number of papers, but this would rule out assessing self-citation trends across many fields over time. We chose to instead analyze the data on a large scale without disambiguated authors and relied on some constraints when exploring the data. This does limit our ability to look at authors across time to see if they even have previous papers to self-cite. Since we cannot track individual authors over time, we cannot control for individual productivity. We cannot determine, therefore, whether men have more papers available to self-cite than women across their careers, which could explain a gap in self-citation rates without any difference in the likelihood of men or women to self-cite.

We looked at all author-to-author citations, where a paper with four authors citing a paper with three authors counts as 12 author-to-author citations, one for each combination. Notice that the fraction of author-to-author self-citations will be less than the fraction of citations that can be considered as self-citations at the paper level. The difference between author-to-author and paper-to-paper self-citations is best illustrated with an example. Suppose that a trio of authors cite another trio of authors with one overlapping member, Colin Jones. The paper Pooja Gupta, Colin Jones, and John Williams (2010) cites the paper Rita Paulson, Colin Jones, and Sarah Erikson (2008). We now have 9 author-to-author pairs (Gupta to Paulson, Gupta to Jones, Gupta to Erikson, Jones to Paulson, etc., etc.) of which one – Colin Jones to Colin Jones – is a self-citation. Thus 1/9-th of the author-to-author citations here is considered self-citation. But if we look at the paper level, the sole citation listed – from Gupta, Jones, and Williams (2010) to Paulson, Jones, and Erikson (2008) is considered a self-citation because Colin Jones is on both papers. Therefore all of the paper-level citations in this case are self-citations.[3]

For the longitudinal analyses, we look at self-citations from the citing year, rather than the cited year. The self-citation rate was calculated for each year independently of other years. For example, the rate for the year 1990 does not look at 1990 and everything before; it only looks at the women's and men's authorships in 1990 and the author-to-author self-citations in 1990.

We calculate the relative rate of men's self-citation to women's self-citation as follows. Let a_w and a_m be the number of women and men authorships respectively. Let s_w and s_m be the number of women's and men's self-citations respectively. Now we answer the question: if we standardize women's self-citation rate to 1, at what rate k to men self-cite? This is calculated by solving the following expression for k:

[3] What if a woman author now has a hyphenated name due to marriage (e.g. Smith-Johnson), but references an article written under her maiden non-hyphenated name (e.g. Smith)? Hyphenated names due to marriage are not of significant concern in our network dataset: there are only 51,270 authorships with hyphens (1.8% of the total), with only a fraction of these likely due to marital name changes.

$$\frac{a_m \times k}{a_w \times 1} = \frac{s_m}{s_w}.$$

RESULTS

In the JSTOR network dataset dating back to 1950, there are 448,389 women authorships and 1,596,125 men authorships.[4] Thus, men represent 78.1% of the authorships for which we know the gender, and women make up the remaining 21.9%.

Within the network dataset, there were 1.7 million papers and 8.2 million citations among the papers. For this analysis, we focused on author-to-author citations. Because papers often have more than one author, there are more author-to-author citations than paper-to-paper citations. For our dataset, we found 39,402,992 author-to-author citations. Of these citations, there are 1,017,362 author-to-author self-citations. Of these, there are 678,768 self-citations by men, 121,923 self-citations by women, and 216,671 self-citations by authors of unknown gender. This means that of self-citations for which we know the gender of the author, men are responsible for 84.8% of the self-citations while women are responsible for 15.2% of the self-citations.

Standardizing women's self-citation rate to 1.0, we solve for men's self-citation rate:

$$\frac{\%\ men\ auth.\ \times men's\ self-cite\ rate\ k}{\%\ women\ auth.\ \times women's\ self-cite\ rate} = \frac{\%\ men's\ self-cites}{\%\ women's\ self-cites}.$$

$$\frac{78.1 \times k}{21.9 \times 1} = \frac{84.8}{15.2}$$

Solving for k, we find that the average man self-cites 56% more often than does the average woman. This is remarkably consistent with Maliniak and colleagues (2013), who analyzed 3000 international relations articles and reported that men authors self-cite 60% more often than do women authors in the same field. Using the JSTOR network dataset, we find that men self-cite their own work 58% more often than women in the field of domestic political science and 68% more often in international political science.

Self-citation patterns by gender

Next we visualize the total number and fraction of self-citations by author gender. We look at absolute numbers rather than the percentage of a paper's citations that are self-citations because there are many papers with one citation that is a self-citation; visualizing the percentage of citations that are self-citations results in long tails and does less to further our understanding.

[4] There were 743,319 authorships for which we could not identify gender. See West et al. (2013) for more details on the method we followed.

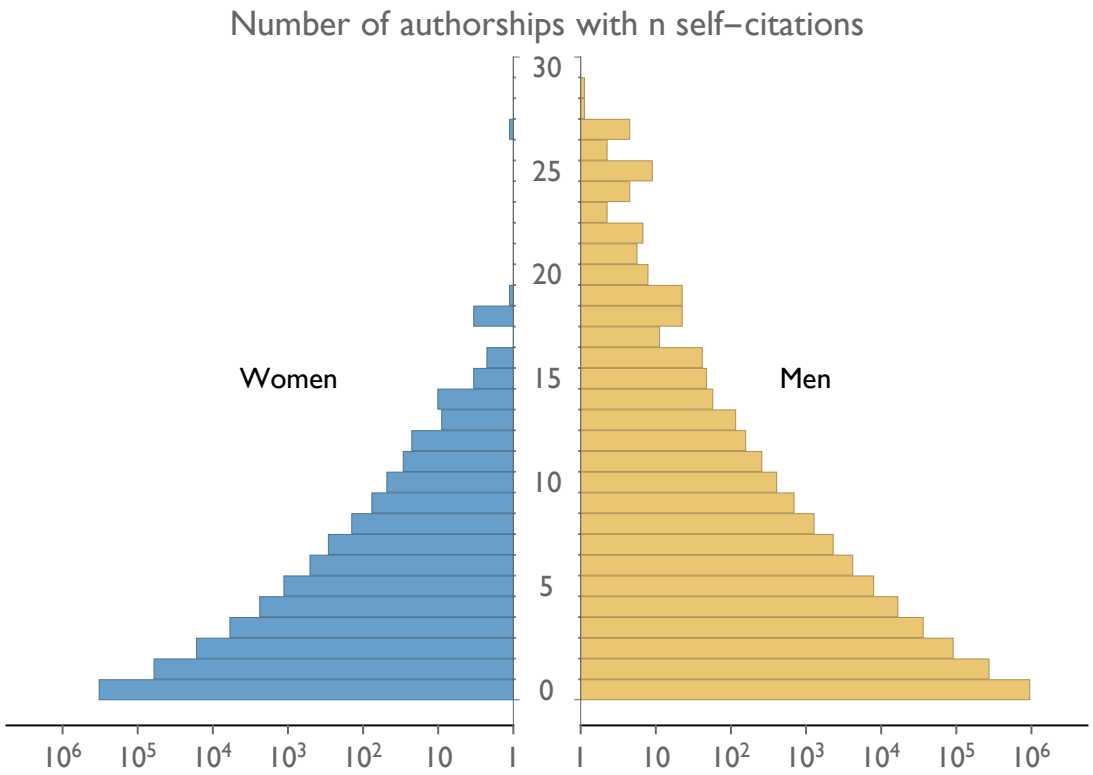

Figure 1. Number of authorship instances where specified number of self-citations occurs, by gender. Each bar's length is equal to the log number of observations of n self-citations.

Men have more frequent counts in all categories of numbers of self-citations, including papers with no self-citations (Figure 1). This is to be expected since men have more instances of authorship overall.

However, **Figure 2** shows us that a greater proportion of women's authorship instances are likely to occur in the zero self-citations category. Women cite themselves one or more times in their papers less often than men do. In other words, compared to men, women are over-represented in the zero self-citations category and underrepresented in terms of citing their papers at all.

Figure 2 shows self-citations grouped by proportions of men's and women's authorships. This shows the rank (1st percentile, 20th percentile, 99th percentile, etc.) in terms of self-citation proportion on the x-axis for men and what rank that same author would have if he were instead part of the women's distribution of authorships on the y-axis. If men and woman behaved similarly in their approaches to self-citation, the corners of the boxes should plot a curve along the x-y diagonal. Instead, wherever there is a difference in the proportion of men and women citing themselves a certain number of times, the boxes deviate from the diagonal.

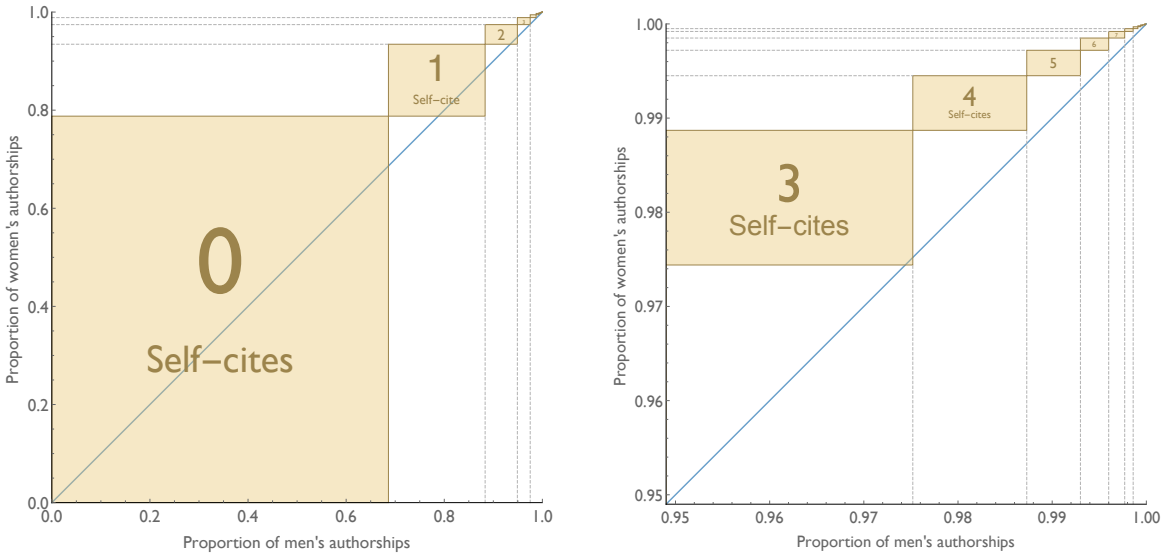

Figure 2. Proportion of authorship instances where specified number of self-citations occurs, by gender. The first half of the figure shows the whole range of possible numbers of self-citations, while the second half zooms in on the area representing 3 self-citations and above. The right edge of each box indicates the proportion of men who cite themselves that number of times, while the upper edge of each box indicates the proportion of women who cite themselves that number of times. The diagonal line represents the point of gender parity, which would bisect the corners of the boxes if the genders behaved identically in patterns of self-citation.

For example, if in a paper you never cite another paper of your own, you are among the vast majority of men (68.6%) and women (78.8%) who also do not cite themselves. If you have one self-citation, you are in the 68th to 88th percentile range for men but the 78th to 93rd percentile for women. With four self-citations, a woman is in the 99th percentile, while a man is in the 98th.

Understanding these distributions is important because they help us see that the gendered nature of self-citation averages is not a result of highly skewed tails. It is the product of the daily activity of the vast majority of academics, those who cite themselves in their papers fewer than five times.

Self-citation rates by year

The ratio of men's self-citation to women's self-citation rose sharply in the 1960s and 1970s, flattening out after about 1980. In the 1950s, the average ratio of the number of men's self-citations relative to women's was 1.27 (with a variance of 0.0199). In the 2000s, this ratio rose to 1.71 (with a variance of 0.0032). This increase is statistically significant (t-test, p-value=1.232e-06). Since 1990, the ratio of men's to women's self-citation is 1.708 (with a variance of 0.00265).

Figure 3 shows the self-citation rate for each year, according to the rate calculated of self-citations from the citing year (rather than the cited year).

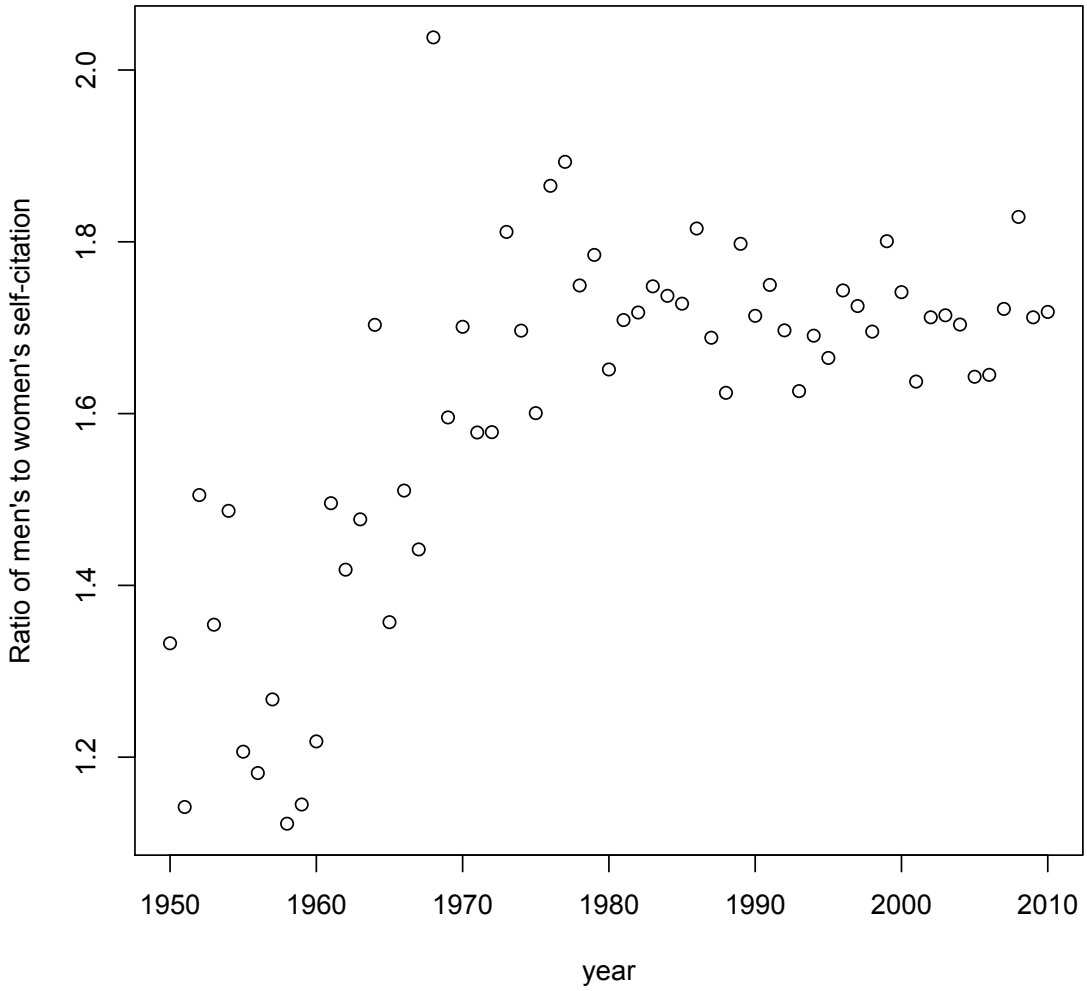

Figure 3. The rate of men's self-citation versus women's self-citation over time. A value of 1.5 means that men cite themselves 50% more than women. See the equation in the methods section for how this rate is calculated. This is based on author-to-author self-citations.

Self-citation rates by field

Though the average ratio shows that men cite their own papers more than women, there is wide variation across fields and subfields. Table 1 shows self-citation rates and gender ratios by fields (as delineated by the hierarchical modeling equation).

Table 1. Self-citation rates for disciplines (i.e., groups at the first level of hierarchical clustering) with at least 5,000 authorships.

Field	Women # self-citations per authorship	Men # self-citations per authorship	Ratio men to women self-citations
Cognitive science	0.25	0.32	1.26
Education	0.22	0.29	1.30
Physical anthropology	0.25	0.33	1.34
Pollution and occupational health	0.25	0.34	1.34
Anthropology	0.20	0.27	1.36
Philosophy	0.21	0.29	1.37
History	0.12	0.16	1.38
Sociology	0.30	0.43	1.43
Ecology and evolution	0.46	0.66	1.44
Organizational and marketing	0.27	0.39	1.45
Classical studies	0.14	0.20	1.45
Law	0.20	0.30	1.47
Veterinary medicine	0.21	0.33	1.52
Radiation damage	0.38	0.57	1.52
Political science-US domestic	0.27	0.43	1.58
Demography	0.28	0.44	1.58
Economics	0.24	0.39	1.65
Political science - international	0.18	0.30	1.68
Molecular & Cell biology	0.30	0.52	1.74
Mathematics	0.20	0.36	1.84
Probability and Statistics	0.24	0.49	2.02

Figure 4 shows men's and women's self-citation patterns by major academic field. Previous research found that women's disadvantage in garnering citations decreased as women made up an increasingly large proportion of the field of economics (Ferber and Brün 2011). We wondered whether there might also be a correlation between the gender composition of a field's authorships and the rate of self-citation. There is no clear relationship between self-citation rate per authorship and the proportion of authorships that are women in a field. The five fields with the lowest women's self-citation rates per authorship (and their corresponding proportions of women authorships in each field) are history (25.6%), classical studies (24.3%), international political science (16.6%), mathematics (7.1%), and general anthropology (37.5%). The fields with the highest women's self-citation rates per authorship are U.S. domestic political science (17.3%), molecular and cell biology (36.5%), sociology (46.1%), radiation damage (30.4%), and ecology and evolution (22.9%).

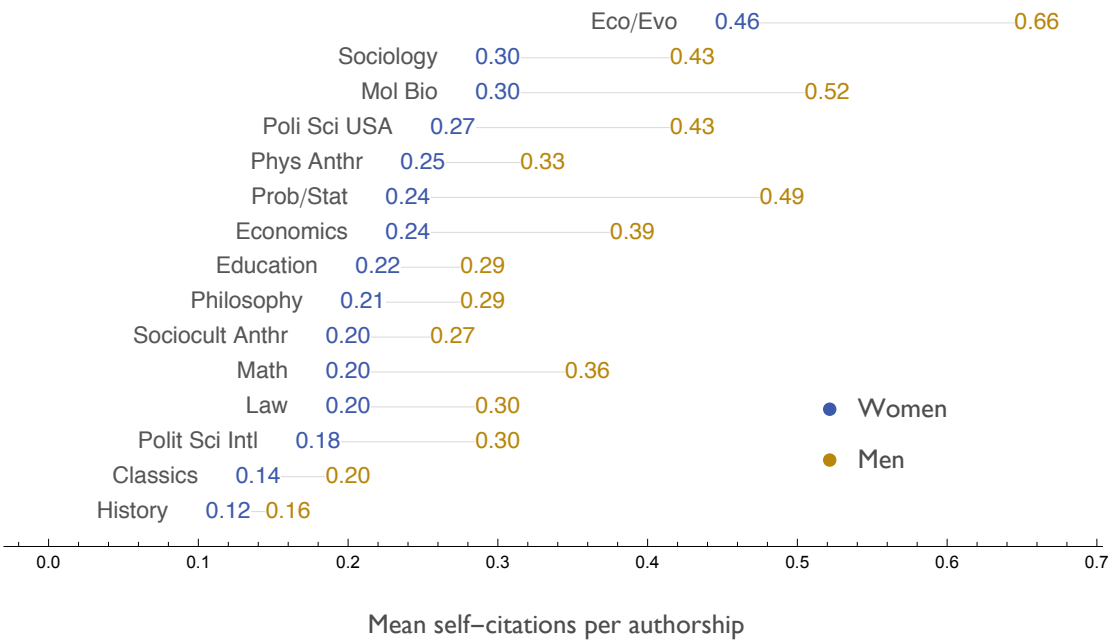

Figure 4. The rate of men's self-citations versus women's self-citations per authorship across major fields. Orange numbers represent men's average number of self-citations per authorship in that field, and blue numbers represent women's average number of self-citations per authorship. This is based on author-to-author self-citations.

Figure 5 shows the relative self-citation rates at the field level. This shows that even within each major academic research field, there is significant variation in the gender ratios of self-citation, depending on the subfield.

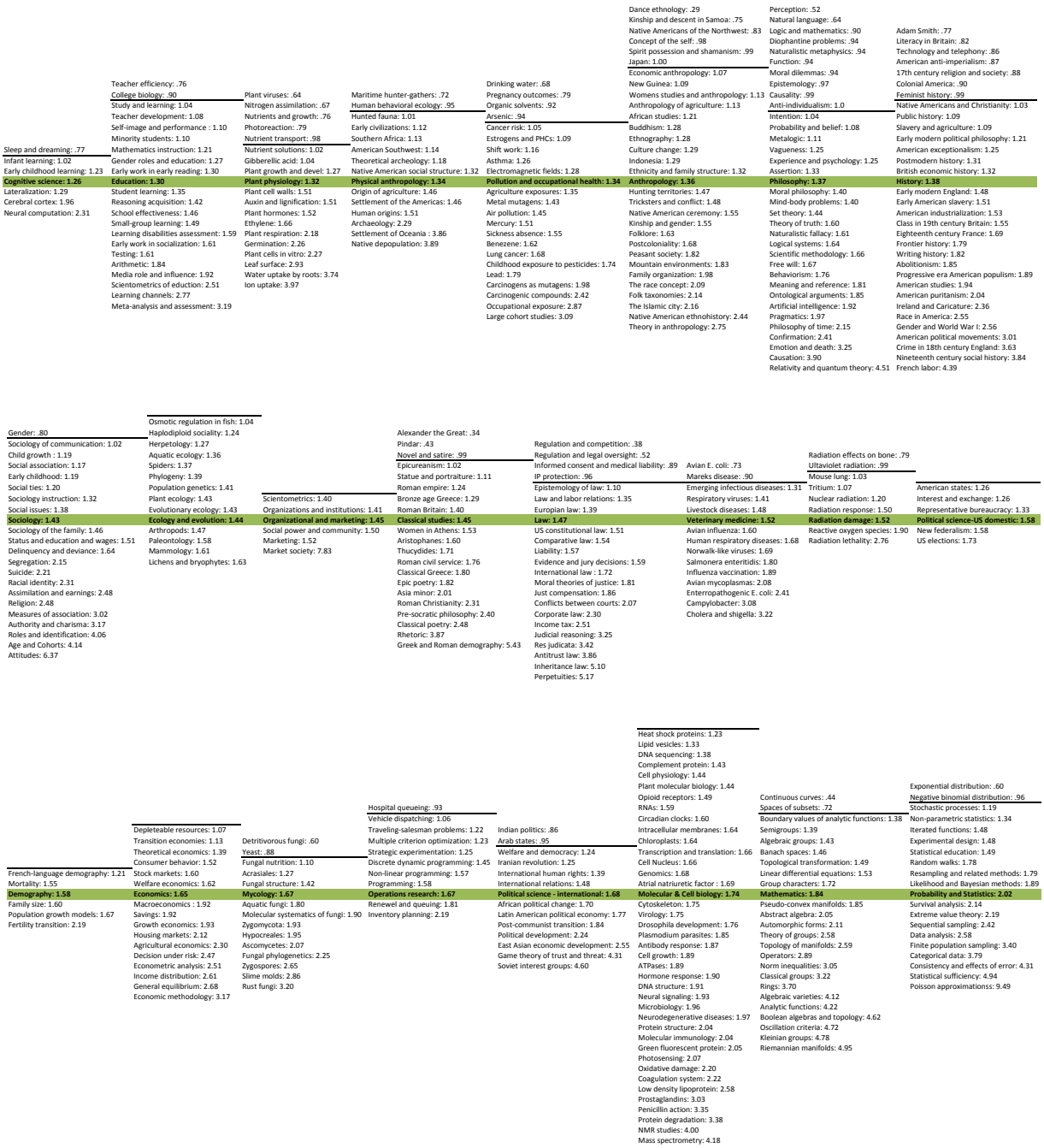

Figure 5. Ratios of men's to women's self-citation rates by field and subfield. Green center line represents self-citation ratio for overall field, with women's self-citation rate set at 1.0. Each subfield is arrayed around its corresponding field based on the subfield's ratio compared to the larger field's ratio. The solid line within each column represents the location of an equal ratio (1.0) of self-citations among men authors and women authors.

How common is self-citation?

To provide more context for the importance of self-citation, we wanted to know: what proportion of citations in an article, on average, are self-citations? This helps to address the relative importance of our gender disparities without disambiguating author names. Within all papers in the JSTOR corpus, 774,113 references were self-citations. Among all 8.2 million references, then, 9.4% are self-citations: references that cite a previous paper authored by one or more of the present paper's author(s). Put another way, across all fields and years about one in 10 references is a self-citation.[5]

Figure 6 presents these results broken down by major academic field. Molecular biology has the highest self-citation rate per reference than other fields, while classical studies has the lowest.

The paper with the most self-citations by its authors is entitled "A Comparison of Whole-Genome Shotgun-Derived Mouse Chromosome 16 and the Human Genome." Published in the journal Science, it includes 220 self-citations to previous papers by the 175 authors. This paper only cites 4 papers that have the same authors on them (representing only 6% of the paper's total 67 references). However, these self-cited papers also have a large number of authors, who are also on the citing paper. Another example is a single-authored paper with 70 self-citations out of 130 references. The paper is titled "Information and the Change in the Paradigm in Economics." Both of these papers show high rates of self-citation per authorship. There are thus different paths to the same end: at one extreme, papers with many authors citing even a few papers with many of the same authors; at the other extreme, sole-authored papers citing many previous papers. These different effects may be differentially likely in different fields.

[5] Within only those papers that included self-citations, there were a total of 3,754,942 references. Among only these papers that cite earlier papers written by their same authors, then, approximately 21% of included references are self-citations!

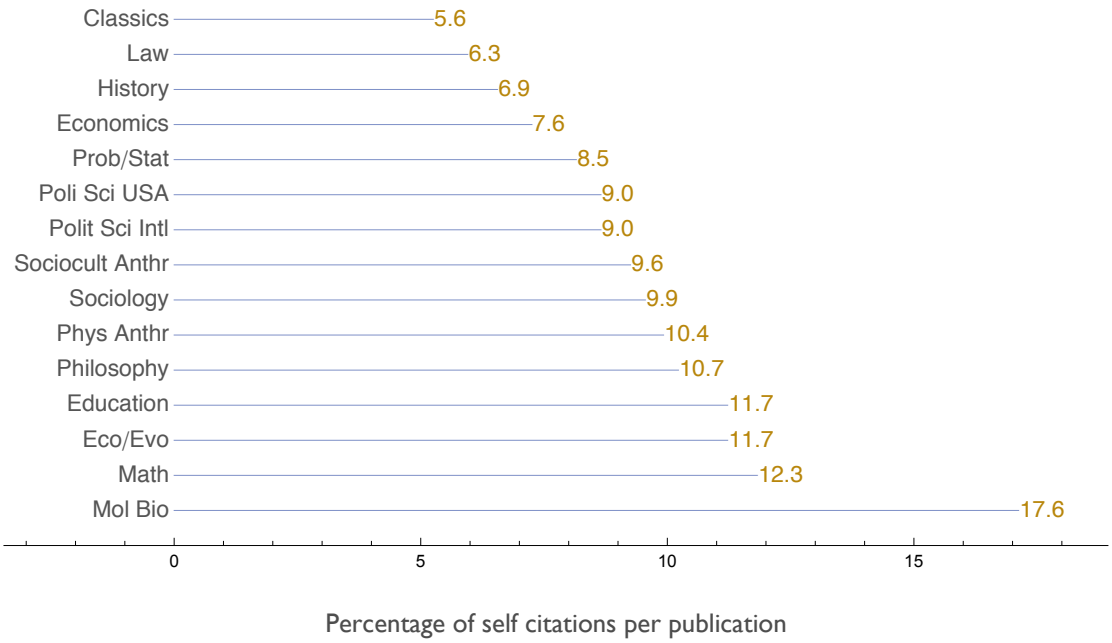

Figure 6. Percent of references that are self-citations per publication across major fields. A value of 10 means 10% of a paper's citations are references to paper(s) written by one or more of the paper's authors.

DISCUSSION

Our study uses an unprecedentedly large dataset of 1.7 million papers across a broad range of academic fields to examine trends in self-citation by academic researchers. Examining 39.4 million author-to-author citations and over 1 million self-citations in this JSTOR database, we uncovered a number of important patterns:

(1) In the last two decades, for every seven self-citations by men, women cited themselves four times (a ratio of 1.7). This ratio rose sharply in the 1960s and 1970s, evening out in the 1980s.

(2) There is wide variation across fields and subfields, and we found no correlation between the proportion of women in a field and women's self-citation rates in that field.

(3) Across the whole JSTOR corpus, about 9.4% of citations are self-citations, indicating that these make up an important fraction of all citations to authors' work.

(4) Men and women differ by more than ten percentage points in how likely they are not to cite themselves in a given paper (68.6% for men vs. 78.8% for women).

We now turn from our findings to speculate on possible mechanisms that might underlie these important trends before finishing with a discussion of their consequential implications.

Possible mechanisms

Why might men academics cite their own previous work more than women academics? While our data include a large number of papers and self-citations, they do not contain variables that allow us to determine the cause of the patterns we identify. However, prior research suggests

several mechanisms that are consistent with our results. We review five possible mechanisms here, which may in some combination contribute to the gender self-citation gap:

(1) Men may self-cite more because they evaluate their abilities more positively than women.
(2) Men face fewer social sanctions for self-promotion.
(3) Men specialize more in academic subfields, and specialization may encourage more self-citation.
(4) Men publish more papers, particularly earlier in their career, and therefore have more work to cite.
(5) Men publish different types of papers, which are the types of papers an academic may be more likely to self-cite.

We will describe the existing evidence for each in turn.

The first two mechanisms – women's lower self-assessments of their accomplishments and greater social sanctions against women who self-promote – are related. Research consistently finds that women evaluate their own abilities more critically, despite evidence of equivalent performance (Correll 2001, 2004) as well as when operating in male-dominated domains (Cech et al. 2011; Thebaud 2010). (Though we did not find that more male-dominated fields result in lower rates of self-citation among women.) Status beliefs about gender shape men's and women's behavior and expectations of themselves and others (Ridgeway 2001, 2014). Because women are perceived as lower status, they are often evaluated more negatively than equally qualified men candidates, by men as well as women (Moss-Racusin et al. 2012). When women seek to actively establish their competence by self-promoting (e.g., by advocating for a raise based on their performance on a project, or asserting their suitability for a leadership position) they often experience backlash from both men and women (Rudman et al. 2012). Though self-promotion enhances competence assessments, it also reduces a woman's likeability (Rudman 1998). These gendered perceptions of self-promotion likely influence perceptions of self-citation, a form of self-promotion in the academic workplace. Status expectations are particularly likely to operate in ambiguous contexts where evaluation criteria are subjective and loosely defined (Fox 2001; Ridgeway 2011), such as those surrounding evaluations of the importance of an academic paper. Historically, women's academic contributions have been undervalued. Rossiter (1993) described the "Matilda Effect," which is the process by which women's scientific ideas are steadily downplayed or ignored and recognition systematically biased in favor of men's intellectual contributions.

Field segregation by gender may also contribute to gender discrepancies in self-citation rates, for two reasons. First, fields have different norms around self-citation. Self-citation rates are higher in the natural sciences (Snyder and Bonzi 1998), which are dominated by men. Second, men tend to specialize more within their academic fields, at least within the disciplines of sociology and linguistics (Leahey 2006); this more specific focus may encourage self-citation. One remaining question for future research is whether specialization might explain gender differences in self-citation tendencies.

In part because men specialize more (Leahey 2006), they tend to produce more total papers per year in most fields (reviewed in Bentley and Adamson 2004), particularly earlier in their careers (Long 1992). Not only does higher productivity lead to more papers for scholars to self-cite;

more productive scholars also generate more highly cited papers (Symonds et al. 2006). The discrepancy in productivity might further exacerbate the gender inequality in citation counts.

Finally, there are also differences in the types of papers produced by men and women; for instance, women are significantly underrepresented as authors of single-authored papers and, on papers with three or more authors, underrepresented in the prestigious positions of first and last author (West et al. 2013). These types of papers may constitute the kind of work that would be in the authors' core areas of research interest, and thus papers they may be more likely to self-cite. It may also be that men publish better papers, which they are then more likely to self-cite. In other words, it may be that those types of papers that women tend to publish disproportionately fewer of are also those that attract more self-citations.

Ratio across time: a widening gap

Another point is that the ratio of men's to women's self-citations begins increasing in the 1960s and 1970s. In recent years, men have cited themselves about 62% more often than women do. There are likely many factors that contribute to such a substantial gap. Though we cannot assess why men began citing themselves at a higher rate than women during this time, two other trends that occurred in this era are worth noting: women began joining the faculty in larger numbers, and scholars began to specialize more. Consequently, there are two possible explanations for the increase in the ratio of men's to women's self-citation rates: (1) scholars in more specialized areas self-cite more and men were increasingly in those areas (as discussed above); and (2) the threat from women faculty moving into the academy led men to self-cite more.

The increasing movement of women into what had previously been almost exclusively "men's jobs" may have led some men to cite their own work more as a way of enhancing their scholarly reputation in the face of growing competition from women. This second possible, albeit conspiratorial-sounding mechanism to explain the *widening gap* – that men may have increased self-citation behavior as a compensatory response as more women entered the academy – has some indirect scholarly evidence. One experimental study tested men's response to group status threat. Men were randomly assigned to a computer-based woman partner who either ascribed to traditional gender roles or reported being a feminist seeking a traditionally male-typed job as a bank manager (the "legitimacy threat" condition). Those assigned to the legitimacy threat condition were more likely to sexually harass the woman – in other words, they were more likely to react to threat in a compensatory way (Maass et al. 2003). Another experiment found that men – but not women – who experienced threats to their gender identity express greater desire for dominance and support for societal dominance hierarchies. When men feel threatened, they compensate with greater displays of masculinity. (Willer et al. 2013). While these studies use different dependent variables than ours, of course, self-promoting is often thought to be associated with masculinity. If self-citation is being used as a form of self-promotion, the increase in men's relative rates as more women entered the academic hierarchy could be explained, at least in part, by group threat felt (subconsciously) by men academics. Again, our data cannot reveal the mechanisms behind this temporal trend; here we present hypotheses bases on the other scholarly literature to encourage future study. Irrespective of the underlying causal mechanisms, we find statistically significant and socially important gender differences in

patterns of self-citation.

Implications

Citation follows a pattern of preferential attachment – the tendency for new citations to refer to papers that are already popular (Fowler and Aksnes 2007; Maliniak et al. 2013). Self-citation increases the number of citations from others (Fowler and Aksnes 2007).
Thus the gender difference in self-citation is likely to be a driver of gender differences in numbers of citations received from other authors. This is not inconsequential: an academic's visibility – reflected in citation counts – has a direct, positive, and significant effect on her salary (Leahey 2007). Citation count is also a key evaluation criterion for hiring and career advancement. Given our finding that nearly 1 in 10 references in a paper is a reference to a paper written by one or more of the current paper's author(s), self-citation is an important contributor to citation counts and academic visibility. Thus, gender discrepancies in self-citation rates have notable consequences for academic careers.

The motives for self-citation vary (Hyland 2003; Safer and Tang 2009; Tang and Safer 2008), but self-citation is one of the few direct ways an academic can increase his or her own citation count. Some scholarly databases (e.g. the Thomson-Reuters Web of Science) provide a separate count of self-citations, while others (e.g., Google Scholar) do not. However, merely encouraging women to cite their own work more often is not a simple solution: it may have unintended consequences due to the backlash against women's self-promotion (Rudman 1998). Furthermore, insisting that scholars self-cite more in order to enhance their reputation could increase irrelevant self-citations. Should this happen, it will become even more difficult to make accurate judgments of the quality and influence of a scholarly work.

Historically, women's academic contributions have been undervalued. Rossiter (1993) described the "Matilda Effect," which is the process by which women's scientific ideas are steadily downplayed or ignored and recognition systematically biased in favor of men's intellectual contributions. When interpreting the impact metrics of scholars' work, university hiring and tenure committees should be aware that women are likely to cite their own work less often. Considering other proposed measures for scientific impact that exclude self-citation (Ferrara and Romero 2013) could make evaluation processes less gender-biased and improve equity in the academic community.

REFERENCES

Bentley, Jerome T., and Rebecca Adamson. 2004. *Gender Differences in the Careers of Academic Scientists and Engineers: A Literature Review*.

Cech, Erin, Brian Rubineau, Susan Silbey, and Carroll Seron. 2011. "Professional Role Confidence and Gendered Persistence in Engineering." *American Sociological Review* 76(5):641–66. Retrieved December 27, 2014 (http://asr.sagepub.com/cgi/doi/10.1177/0003122411420815).

Correll, Shelley J. 2001. "Gender and the Career Choice Process: The Role of Biased Self-Assessments." *American Journal of Sociology* 106(6):1691–1730.

Correll, Shelley J. 2004. "Constraints into Preferences : Gender, Status, and Emerging Career Aspirations." *American Sociological Review* 69(1):93–113.

Ferber, Marianne A., and Michael Brün. 2011. "The Gender Gap in Citations: Does It Persist?" *Feminist Economics* 17(1):151–58. Retrieved October 25, 2012 (http://www.tandfonline.com/doi/abs/10.1080/13545701.2010.541857).

Ferrara, Emilio, and Alfonso E. Romero. 2013. "Scientific Impact Evaluation and the Effect of Self-Citations: Mitigating the Bias by Discounting the H-Index." *Journal of the American Society for Information Science and Technology* 64(11):2332–39. Retrieved (http://doi.wiley.com/10.1002/asi.22976).

Fowler, James H., and Dag W. Aksnes. 2007. "Does Self-Citation Pay?" *Scientometrics* 72(3):427–37. Retrieved January 22, 2014 (http://link.springer.com/10.1007/s11192-007-1777-2).

Fox, Mary Frank. 2001. "Women, Science, and Academia: Graduate Education and Careers." *Gender & Society* 1(5):654–66.

Hutson, Scott R. 2006. "Self-Citation in Archaeology: Age, Gender, Prestige, and the Self." *Journal of Archaeological Method and Theory* 13(1):1–18. Retrieved March 7, 2014 (http://link.springer.com/10.1007/s10816-006-9001-5).

Hyland, Ken. 2003. "Self-Citation and Self-Reference: Credibility and Promotion in Academic Publication." *Journal of the American Society for Information*

Science and Technology 54(3):251–59. Retrieved (http://doi.wiley.com/10.1002/asi.10204).

Larivière, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. "Global Gender Disparities in Science." *Nature* 504:211–13.

Leahey, Erin. 2006. "Gender Differences in Productivity: Research Specialization as a Missing Link." *Gender & Society* 20(6):754–80. Retrieved February 17, 2011 (http://gas.sagepub.com/cgi/doi/10.1177/0891243206293030).

Leahey, Erin. 2007. "Not by Productivity Alone: How Visibility and Specialization Contribute to Academic Earnings." *American Sociological Review* 72(4):533–61. Retrieved March 5, 2011 (http://asr.sagepub.com/cgi/doi/10.1177/000312240707200403).

Lieberson, Stanley, Susan Dumais, and Shyon Baumann. 2000. "The Instability of Androgynous Names: The Symbolic Maintenance of Gender Boundaries." *American Journal of Sociology* 105(5):1249–87. Retrieved (http://www.journals.uchicago.edu/doi/abs/10.1086/210431).

Long, J. Scoit. 1992. "Measures of Sex Differences in Scientific Productivity." *Social Forces* 71(1):159–78. Retrieved (http://sf.oxfordjournals.org/content/71/1/159.short).

Maass, Anne, Mara Cadinu, Gaia Guarnieri, and Annalisa Grasselli. 2003. "Sexual Harassment under Social Identity Threat: The Computer Harassment Paradigm." *Journal of personality and social psychology* 85(5):853–70.

Maliniak, Daniel, Ryan Powers, and Barbara F. Walter. 2013. *The Gender Citation Gap in International Relations*. Retrieved November 9, 2013 (http://www.journals.cambridge.org/abstract_S0020818313000209).

Moss-Racusin, Corinne A., John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2012. "Science Faculty's Subtle Gender Biases Favor Male Students." *Proceedings of the National Academy of Sciences of the United States of America*. Retrieved October 3, 2012 (http://www.ncbi.nlm.nih.gov/pubmed/22988126).

Ridgeway, C. L. 2014. "Why Status Matters for Inequality." *American Sociological Review* 79(1):1–16. Retrieved December 23, 2013 (http://asr.sagepub.com/cgi/doi/10.1177/0003122413515997).

Ridgeway, Cecilia L. 2001. "Gender, Status, and Leadership." *Journal of Social Issues* 57(4):637–55.

Ridgeway, Cecilia L. 2011. *Framed by Gender: How Gender Inequality Persists in the Modern World*. Oxford University Press.

Rossiter, Margaret W. 1993. "The Matthew Matilda Effect in Science." *Social Studies of Science* 23(2):325–41. Retrieved July 3, 2013 (http://sss.sagepub.com/cgi/doi/10.1177/030631293023002004).

Rosvall, Martin, and Carl T. Bergstrom. 2011. "Multilevel Compression of Random Walks on Networks Reveals Hierarchical Organization in Large Integrated Systems." *PloS one* 6(4):e18209. Retrieved October 31, 2012 (http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3072965&tool=pmcentrez&rendertype=abstract).

Rudman, Laurie A. 1998. "Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management." *Journal of Personality and Social Psychology* 74(3):629–45. Retrieved (http://www.ncbi.nlm.nih.gov/pubmed/9523410).

Rudman, Laurie A., Corinne A. Moss-Racusin, Julie E. Phelan, and Sanne Nauts. 2012. "Status Incongruity and Backlash Effects: Defending the Gender Hierarchy Motivates Prejudice against Female Leaders." *Journal of Experimental Social Psychology* 48(1):165–79. Retrieved March 4, 2013 (http://linkinghub.elsevier.com/retrieve/pii/S0022103111002514).

Safer, Martin a., and Rong Tang. 2009. "The Psychology of Referencing in Psychology Journal Articles." *Perspectives on Psychological Science* 4(1):51–53. Retrieved January 21, 2014 (http://pps.sagepub.com/lookup/doi/10.1111/j.1745-6924.2009.01104.x).

Snyder, H., and S. Bonzi. 1998. "Patterns of Self-Citation across Disciplines (1980-1989)." *Journal of Information Science* 24(6):431–35. Retrieved March 8, 2014 (http://jis.sagepub.com/cgi/doi/10.1177/016555159802400606).

Symonds, Matthew R. E., Neil J. Gemmell, Tamsin L. Braisher, Kylie L. Gorringe, and Mark a Elgar. 2006. "Gender Differences in Publication Output: Towards an Unbiased Metric of Research Performance." *PloS one* 1(1):e127. Retrieved October 25, 2012

(http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1762413&tool=pmcentrez&rendertype=abstract).

Tang, Rong, and Martin a. Safer. 2008. "Author-Rated Importance of Cited References in Biology and Psychology Publications." *Journal of Documentation* 64(2):246–72. Retrieved February 8, 2014 (http://www.emeraldinsight.com/10.1108/00220410810858047).

Thebaud, S. 2010. "Gender and Entrepreneurship as a Career Choice: Do Self-Assessments of Ability Matter?" *Social Psychology Quarterly* 73(3):288–304. Retrieved December 28, 2014 (http://spq.sagepub.com/cgi/doi/10.1177/0190272510377882).

West, Jevin D., Jennifer Jacquet, Molly M. King, Shelley J. Correll, and Carl T. Bergstrom. 2013. "The Role of Gender in Scholarly Authorship" edited by Lilach Hadany. *PloS one* 8(7):e66212. Retrieved July 26, 2013 (http://arxiv.org/abs/1211.1759).

Willer, Robb, Christabel L. Rogalin, Bridget Conlon, and Michael T. Wojnowicz. 2013. "Overdoing Gender: A Test of the Masculine Overcompensation Thesis." *American Journal of Sociology* 118(4):980–1022.