Gender Differences in Synchronous and Diachronous Self-citations

Gita Ghiasi[*], Vincent Larivière[**], Cassidy R. Sugimoto[***]

[*]g_ghia@encs.concordia.ca
Department of Mechanical and Industrial Engineering, Concordia University, 1455 De Maisonneuve Blvd. W., Montréal, QC, H3G 1M8, Canada
**vincent.lariviere@umontreal.ca
École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, QC, H3C 3J7, Canada
*** sugimoto@indiana.edu
School of Informatics and Computing, Indiana University, 1320 E. 10th St, Bloomington, IN, 47405, United States of America

ABSTRACT

Citation rates are increasingly used as a currency of science, providing a basis to reward a scientist. Self-citations, an inevitable part of scholarly communication, may contribute to the inflation of citation counts and impose a considerable impact on research evaluation and academic career advancements. Self-citations are classified into two types in this study: synchronous self-citations (self-citations an author gives) and diachronous self-citations (self-citations an author receives). The main objective of this paper is to provide a comprehensive gendered analysis of synchronous and diachronous self-citations across all scientific disciplines. For this purpose, citation data of 12,725,171 articles published in 2008-2014 are extracted from Web of Science and are further scrutinized for articles of each gender. The findings reveal that men receive citations from their own papers at a higher rate than their women counterparts. They also tend to give more citations to their own publications. Gender gap in citation impact decreases when first-author's diachronous citations are eliminated in the impact analysis. However, the gap does not vary when all-authors' diachronous citations are excluded. The results of this research is important for effective gender-related policy-making in the science and technology arena.

INTRODUCTION

Women have been long susceptible to the "Matilda effect" in science (Rossiter, 1993)—the opposite of so-called "Matthew Effect"(Merton, 1968), an allusion to a well-known colloquialism "the rich get richer and the poor get poorer"—where the provisions of more credit to eminent scientists afflict women scientists by systematically repressing or ignoring their contribution to research and attributing their work to their male colleagues.

This means that the scientific discoveries of women get little to no credit for the same quality work as their male peers, only because of their gender. This consequence can expose a more prominent scientist—generally a male scientist—to ever-more resources and thereby more recognition and credit (for a similar work) than their contemporary female peers, leaving female scientists of the field unknown and invisible (Duran & Lopez, 2014).

As a result, women's contribution to science is often overlooked in the receipt of prestigious research award and grants (Lincoln, Pincus, Koster, & Leboy, 2012) and their research is less valued (Trix & Psenka, 2003; Wennerås & Wold, 1997). Women are commonly associated with low-quality publications and are subject of lower collaboration interest (Knobloch-Westerwick, Glynn, & Huge, 2013). Their publications are cited less frequently than their male peers after controlling for authorship positions (i.e. sole, first and last author) (Larivière, Ni, Gingras, Cronin, & Sugimoto, 2013), and affiliation, tenure status, methodology and context (Maliniak, Powers, & Walter, 2013).

Matilda effect has repeatedly shown to be present at citation level: although being published in journals with higher citation rates (Impact Factor), female-authored papers receive lower number of citations (Ghiasi, Larivière, & Sugimoto, 2015; Knobloch-Westerwick & Glynn, 2013; Larivière, 2014). Yet, citation rates are increasingly used as a currency of science—which mirror a base from which to reward a scientist (Merton, 1973)—and have become lamentably popular as the determinants of hiring, reappointment, tenure, promotion (Holden, Rosenberg, & Barker, 2005) and faculty salary (Toutkoushian, 1994), disfavoring women in their scientific research system.

Along these lines, self-citations may not only help inflate an author's citation counts, but impose a considerable impact on scholarly careers of academic researchers—one additional self-citation to a given paper attracts one extra citation from other researchers after one year and three extra citations after five years (Fowler & Aksnes, 2007). As citation rates are becoming a popular measure for research evaluations, self-citations are increasingly being used as a convenient means for manipulating the rewarding system of a scientist—paving the way for a researcher to gain more recognition and thus become more visible and influential.

Self-citations are the inevitable consequence of expanding on earlier study or furthering research in a specific field, but are also served as a tactic of manipulation for increasing a researcher's h-index, which is proved to have a positive impact on the academic ranking (Bartneck & Kokkelmans, 2010) of a scientist. Hence, self-citations might add to the persistence of Matilda effect in science and play a major role in forming the scientific system that disfavors women in hiring and tenure procedures, salary decisions, and workplace advancements.

Gendered analysis of self-citations is very nascent and is limited to (Hutson, 2006; King, Correll, Jacquet, Bergstrom, & West, 2015; Susarla, Swanson, Lopez, Peacock, & Dodson, 2015). Among these studies, no significant effect between gender and self-citations have found (Hutson, 2006; Susarla et al., 2015). However, Hutson (2006) further scrutinized this finding and found that men cite themselves more often than women cite themselves (even when considering the rate of self-citations within the text of a paper). This is in line with the work of King et al. (2015) who also found an increasing gender gap in self-citations over the last 50 years.

Self-citations are generally categorized into two types (Aksnes, 2003), namely synchronous and diachronous self-citations. The former applies when the author cites his/her previous paper(s) in the paper that is being studied and the latter is when the paper that is being studied is cited by the author in one of his/her subsequent papers. Within the relevant literature, there is a gap in differentiating between these two types of self-citations and analyzing two statistics: cited by (or to) self and cited by (or to) all authors of the paper. This study tries to fill this gap and provides a comprehensive gendered analysis of synchronous and diachronous self-citations across all scientific fields.

METHOD

Article data is gathered from Thomson Reuters' Web of Science (WoS)—a comprehensive database of peer-reviewed publications and citations, which contains the Science Citation Index Expanded, the Social Science Citation Index and the Arts and Humanities Citation Index. The classification of scientific disciplines are based on that of the U.S. National Science Foundation's (NSF), which assigns each journal to only one discipline and one specific specialty and is necessary to avoid multiple counting of articles that are published in multidisciplinary journals. NSF classifies disciplines into 14 categories, namely Arts, Biology, Biomedical Research, Chemistry, Clinical Medicine, Earth and Space, Engineering and Technology, Health, Humanities, Mathematics, Physics, Professional Fields, Psychology, and Social Sciences.

A total of 12,725,171 articles are identified for the years 2008-2014 and are assigned to the aforementioned disciplines (Table 1). The focus of this study is on articles published after the year 2008, because WoS covers the full first names of authors from the year 2008 (i.e. essential for assigning gender to the authors).

Table 1. Number of articles per discipline over the period 2008-2014

Discipline	Number of Articles	Discipline	Number of Articles
Arts	142,015	Health	319,803
Biology	800,856	Humanities	574,585
Biomedical Research	1,365,778	Mathematics	324,398
Chemistry	1,059,575	Physics	885,975
Clinical Medicine	4,242,232	Professional Fields	348,092
Earth and Space	555,983	Psychology	262,391
Engineering and Technology	1,366,841	Social Sciences	476,647

Gender of WoS authors is further assigned by matching authors' given names with universal and country-specific existing name and gender databases, including U.S. Census, WikiName, Wikipedia, France and Quebec lists, and country-specific lists (which is explained in detail in (Larivière et al., 2013)).

Citation data is normalized for publication year and subject area, and is measured as the average yearly number of citations to a given paper from its year of publication to end of the year 2014, divided by the average yearly number of citations received by all papers published in the same year and in the same field.

Diachronous self-citations are identified by matching author names of citing articles with author names on a cited article (authors with the same last name and first name abbreviations). Citation data is then normalized, excluding diachronous self-citations received by both first author's and co-authors' papers. Finally, gender gap is calculated as the difference between average citation rates of articles first-authored by men and women relative to average citation rates of male first-authored articles.

Similarly, synchronous self-citations are identified where articles listed as references of a given paper have authors with similar name to authors of that paper (same last name and first name abbreviations), and are further grouped into two measures: citations to first-author's articles and citations to articles written by any of the authors. Self-referencing rate is defined for each gender as the ratio of self-citations to total number of references and gender gap in self-referencing rate is then calculated. The smaller the percentage, the more equal researchers of each gender self-cite.

RESULTS

Citation impact of articles first-authored by male scientists is higher in all different disciplines (except Arts). Gender gap in citation impact is 9% for all scientific disciplines (Fig. 1) and is the highest in Biomedical Research, Psychology and Chemistry (14%). It is the lowest in Engineering and Technology and Biology (3%) and is non-existent in Arts (female-authored papers in arts receive more citations). However, the gender gap in citation impact slightly shrinks when excluding the first-author's diachronous self-citations, revealing more equal rate of citations received by articles of each gender. Interestingly, when eliminating all-authors' diachronous self-citations, the same rate for gender gap is reached as if all the self-citations are included. This applies to all scientific disciplines in all the years (Fig. 2) and might imply that male first-authored articles tend to receive citations from their subsequent publications at a higher rate than female first-authored articles. Nevertheless, when women are first authors, their papers might receive citations from their co-authors at a higher rate than that of their male counterparts—which, in part, explains the equal percentages for gender gap including and excluding all author's self-citations.

Figure 1: Gender gap in citation impact of scientific papers by discipline, including and excluding diachronous self-citations (2008-2014)

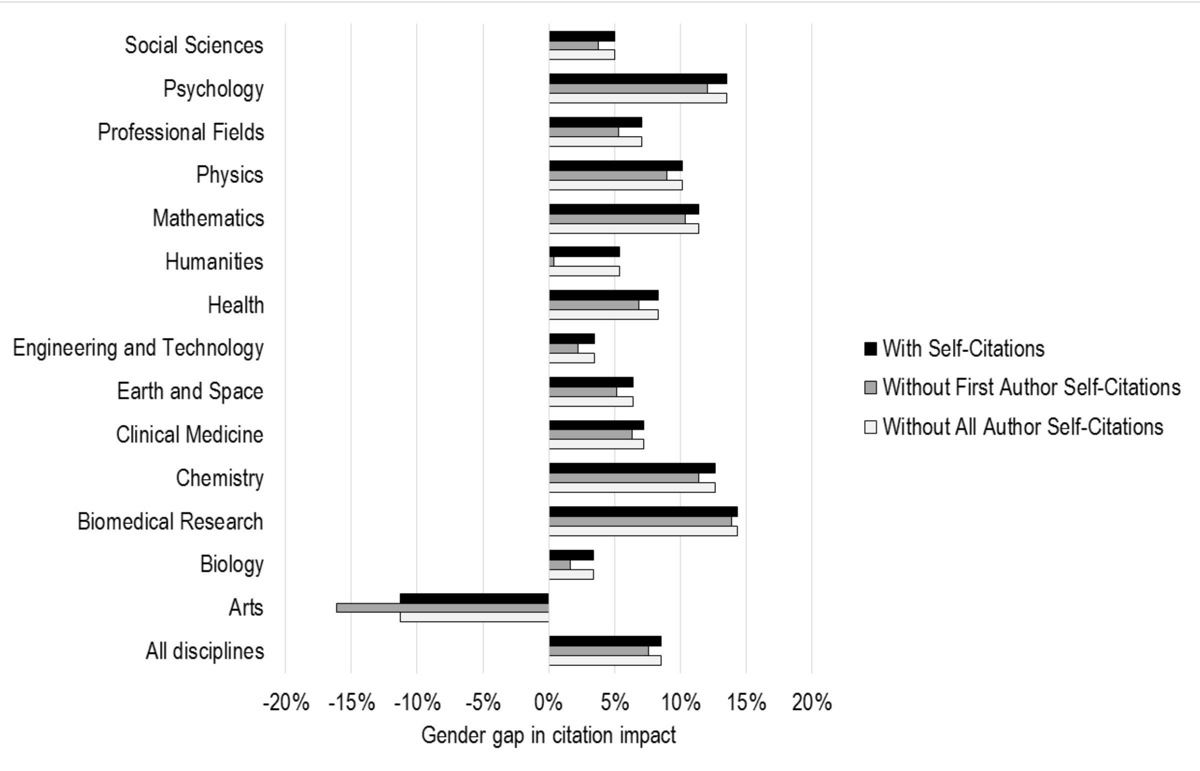
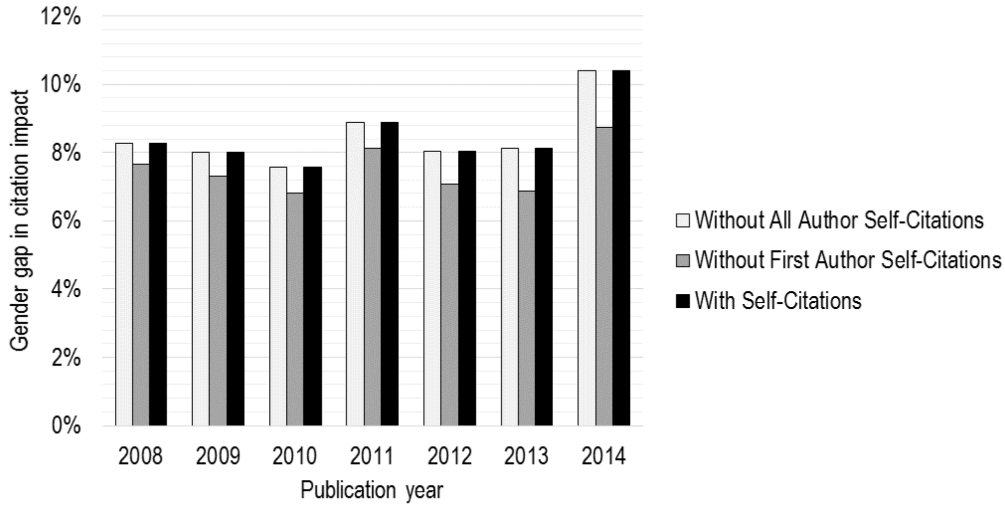


Figure 2: Gender gap in citation impact of scientific papers by year, including and excluding diachronous self-citations



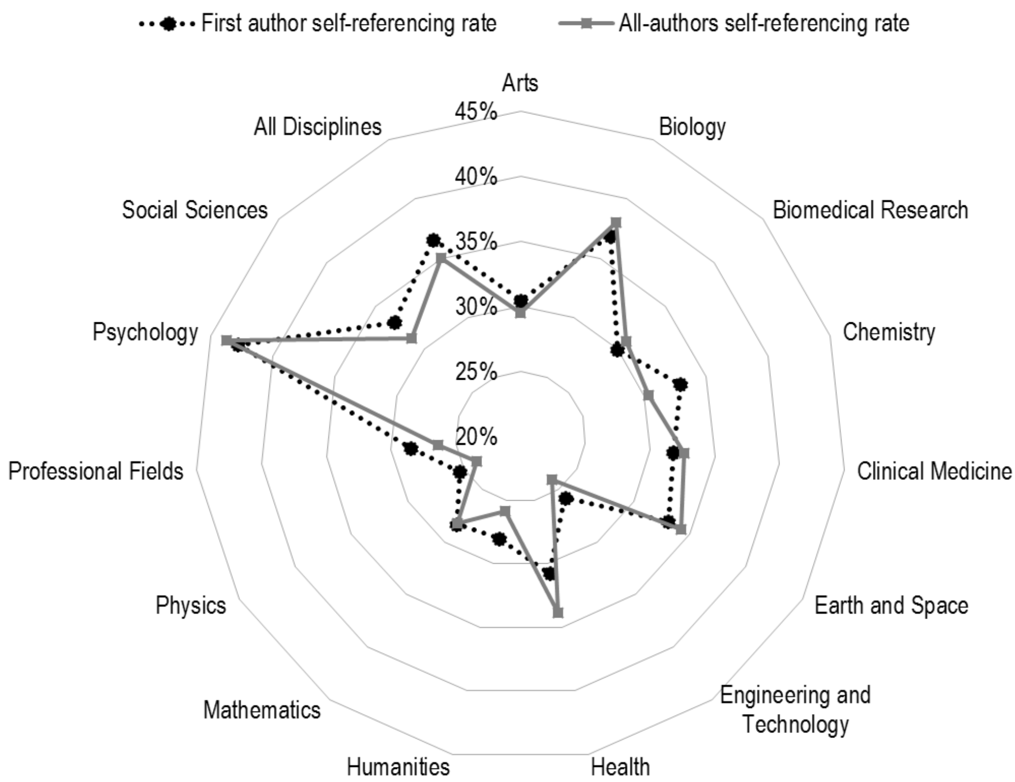
The analysis of synchronous self-citation rates (aka, self-referencing rates) shows that men, when listed as the first author of a paper, tend to cite their past works or any of the authors' previous works at a higher rate than their female peers across all the disciplines (Table 2). The highest rate of self-referencing (for both female and male first-authored articles) is in Mathematics—which might be due to the fact that research in this particular field is expanding on earlier hypotheses and methods—and the lowest rate is in Biomedical Research and Clinical Medicine, showing that new and dynamic discoveries are furthering research in these areas.

Table 2. First author and all-authors synchronous self-citation (self-referencing) rate for female and male authored papers by discipline (2008-2014)

	First-author self-referencing rate		All-author self-referencing rate	
Discipline	**Female**	**Male**	**Female**	**Male**
Arts	3.3%	4.7%	4.6%	6.5%
Biology	2.1%	3.4%	3.5%	5.7%
Biomedical Research	1.5%	2.2%	3.1%	4.5%
Chemistry	2.3%	3.5%	4.0%	5.7%
Clinical Medicine	1.5%	2.3%	3.0%	4.5%
Earth and Space	2.5%	3.8%	4.3%	6.6%
Engineering and Technology	2.9%	3.8%	4.5%	5.9%
Health	2.2%	3.1%	3.4%	5.1%
Humanities	3.8%	5.2%	4.9%	6.6%
Mathematics	4.9%	6.9%	6.6%	9.2%
Physics	3.4%	4.5%	5.5%	7.3%
Psychology	2.3%	4.1%	3.6%	6.5%
Social Sciences	2.4%	3.6%	3.4%	5.0%
Professional Fields (Others)	1.8%	2.6%	2.7%	3.7%
All Disciplines	1.9%	3.1%	3.5%	5.3%

Gender gap in first-author self-referencing rate is 37% and in all-author self-referencing rate is 35% (Fig. 3). The lowest gender gap in self-referencing rate is in Physics and Engineering and largest is in Psychology. Surprisingly, psychology is among the fields with highest rate of female authorship (women represent ~50% of total authorship), whereas physics and engineering are the most male-dominated disciplines (women account for only 20% of total authorship) (Larivière, 2014). The lower gender gap in self-referencing rate might be associated with the selection effect in male-dominated fields—the fact that women need to be highly competent in order to survive or stay in in the most male-dominated fields—and their research, hence, serves as an important resource upon which new discoveries are grounded.

Figure 3: Gender gap in self-referencing rate by discipline (2008-2014)



DISCUSSION

Articles that are authored by a male author, tend to receive higher rate of citations from his subsequent publications. When first-author diachronous citations are removed from the citation impact analysis, the gender gap decreases. Nevertheless, the gender gap in citation impact does not vary when all-author diachronous citations are excluded. This shows that a paper that is authored by a women, might receive higher citation rates from her co-authors' publications than a paper authored by a man, which might be associated to gender differences in self-promotions: although women self-promote their own works at lower level and their publications receive lower recognition (citations) from the scientific community than those of their male peers, their work is promoted and recognized at higher rate by their immediate co-authors.

Narrowing the focus to synchronous self-citations or (self-references), it can be noted that men, in their papers, refer to their previous works at a higher rate than women scientists. However, the gender gap in self-referencing rate is the lowest in Physics and Engineering—the most male-dominated disciplines—which is bound to the exceptional competence and expertise of women researchers in the most male dominated fields (Dryburgh, 1999; Ghiasi et al., 2015), where their work serve as a basis from which new studies are made.

The exposure of women scientists to Matilda effect shed light on the lower citation impact of their work (although being published in higher Impact Factor journals) (Larivière, 2014). One of the elucidations for gender differences in citations is that men receive citations from their own papers at a higher rate than their women counterparts and they tend to give more citations to their own publications. Citation patterns of an article have proved to fall under the Matthew effect in the sense that papers with high number of citations continue to be cited at a higher rate (Merton, 1968, 1988). Therefore, gender differences in self-citations play a major role in attracting more citations, have a direct impact on the h-index score of an author, and might contribute to gender inequality in evaluation, hiring, promotion and pay in academia.

The results of this research is thus of utmost importance for effective policy-making, redefining scientific reward and evaluation system with the use of gender equity measures. Recognition of women's contributions to scientific research can help identify mismatches in science and technology policies that can thwart gender parities in scientific performance, which subsequently gear toward a more equitable society.

REFERENCES

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, *12*(3), 159–170.

Bartneck, C., & Kokkelmans, S. (2010). Detecting h-index manipulation through self-citation analysis. *Scientometrics*, *87*(1), 85–98.

Dryburgh, H. (1999). WORK HARD, PLAY HARD Women and Professionalization in Engineering—Adapting to the Culture. *Gender & Society*, *13*(5), 664–682.

Duran, A., & Lopez, D. (2014). Women from Diverse Backgrounds in the Science, Technology, Engineering, and Math (STEM) Professions: Retention and Career Development. *Impact of Diversity on Organization and Career Development*, 214.

Fowler, J. H., & Aksnes, D. W. (2007). Does self-citation pay? *Scientometrics*, *72*(3), 427–437.

Ghiasi, G., Larivière, V., & Sugimoto, C. R. (2015). On the Compliance of Women Engineers with a Gendered Scientific System. *PloS one*, *10*(12), e0145931.

Holden, G., Rosenberg, G., & Barker, K. (2005). Bibliometrics: A potential decision making aid in hiring, reappointment, tenure and promotion decisions. *Social work in health care*, *41*(3-4), 67–92.

Hutson, S. R. (2006). Self-Citation in Archaeology: Age, Gender, Prestige, and the Self. *Journal of Archaeological Method and Theory*, *13*(1), 1–18.

King, M. M., Correll, S. J., Jacquet, J., Bergstrom, C. T., & West, J. D. (2015). *Men set their own cites high: Gender and self-citation across fields and over time*. Working paper. Retrieved from http://www. eigenfactor. org/gender/self-citation/SelfCitation. pdf.

Knobloch-Westerwick, S., & Glynn, C. J. (2013). The Matilda Effect-Role Congruity Effects on Scholarly Communication: A Citation Analysis of Communication Research and Journal of Communication Articles. *Communication Research*, *40*(1), 3–26.

Knobloch-Westerwick, S., Glynn, C. J., & Huge, M. (2013). The Matilda Effect in Science Communication: An Experiment on Gender Bias in Publication Quality Perceptions and Collaboration Interest. *Science Communication*, *35*(5), 603–625.

Larivière. (2014, March). Femmes et sciences : les premières données mondiales valident l'inégalité | Acfas | magazine Découvrir | mars 2014. *Acfas.ca*. Retrieved October 26, 2014, from http://www.acfas.ca/publications/decouvrir/2014/03/femmes-sciences-premieres-donnees-mondiales-valident-l-inegalite

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, *504*(7479), 211–213.

Lincoln, A. E., Pincus, S., Koster, J. B., & Leboy, P. S. (2012). The Matilda Effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science*, *42*(2), 307–320.

Maliniak, D., Powers, R., & Walter, B. F. (2013). The gender citation gap in international relations. *International Organization*, *67*(04), 889–922.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56–63.

Merton, R. K. (1973). The normative structure of science. *The sociology of science: Theoretical and empirical investigations*, *267*.

Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 606–623.

Rossiter, M. W. (1993). The Matthew Matilda effect in science. *Social studies of science*, *23*(2), 325–341.

Susarla, S. M., Swanson, E. W., Lopez, J., Peacock, Z. S., & Dodson, T. B. (2015). Does Self-Citation Influence Quantitative Measures of Research Productivity Among Academic Oral and Maxillofacial Surgeons? *Journal of oral and maxillofacial surgery : official journal of the American Association of Oral and Maxillofacial Surgeons*, *73*(10), 1981.e1–7.

Toutkoushian, R. K. (1994). Using Citations to Measure Sex Discrimination in Faculty Salaries. *The Review of Higher Education*, *18*(1), 61–82.

Trix, F., & Psenka, C. (2003). Exploring the color of glass: Letters of recommendation for female and male medical faculty. *Discourse & Society*, *14*(2), 191–220.

Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, *387*(6631), 341–343.