

# Who acknowledges who? A gender analysis

Adèle Paul-Hus<sup>1</sup>, Philippe Mongeon<sup>2</sup>, Maxime Sainte-Marie<sup>1</sup> and Vincent Larivière<sup>3</sup>

<sup>1</sup> [adele.paul-hus@umontreal.ca](mailto:adele.paul-hus@umontreal.ca) ; [maxime.sainte-marie@umontreal.ca](mailto:maxime.sainte-marie@umontreal.ca)

École de bibliothéconomie et des sciences de l'information, Université de Montréal, PO Box 6128, Downtown Station, Montreal, Quebec, H3C 3J7 (Canada)

<sup>2</sup> [philippe.mongeon@ps.au.dk](mailto:philippe.mongeon@ps.au.dk)

Danish Centre for Studies in Research and Research Policy, Department of Political Science, Aarhus University, Aarhus, Denmark

<sup>3</sup> [vincent.lariviere@umontreal.ca](mailto:vincent.lariviere@umontreal.ca)

École de bibliothéconomie et des sciences de l'information, Université de Montréal, PO Box 6128, Downtown Station, Montreal, Quebec, H3C 3J7 (Canada) and Université du Québec à Montréal, Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Observatoire des Sciences et des Technologies (OST), CP 8888, Succ. Centre-Ville, H3C 3P8 Montréal, Qc. (Canada)

## Abstract

Acknowledgements found in scholarly papers allow for credit attribution among non-authors—from individuals to organizations—that contributed to a piece of research. As such, they are associated with a different kind of recognition than authorship. While several studies have shown that social factors affect authorship and citation practices, few analyses have been performed on acknowledgements. Based on 878,250 addressees mentioned in 291,167 papers published in 2015 retrieved from Web of Science, the objective of this work-in-progress is to better understand such credit attribution practices and how gender may influence them. Our results show that gender disparities generally found in authorship can be extended to acknowledgements, and that women are even more underrepresented in the acknowledgement than in the authors' list of a paper. Our findings also confirm that women acknowledge proportionally more women. These results suggest that gender plays a role in the entire spectrum of credit attribution practices, from authorships to acknowledgements.

## Introduction

Acknowledgements found in scientific papers are a public testimony of authors' gratitude and recognition that can help reveal contributions of varied nature made by individuals, institutions and organizations. The rich information acknowledgements convey allows credit distribution among authors and other contributors. However, in the reward system of science (Merton, 1973) where authorship constitutes the main means to accumulate "symbolic capital" (Bourdieu, 1975), a mention in the acknowledgement is not associated to the same kind of recognition as authorship. Given the hierarchical structure of the scientific community, credit attribution can be difficult to disentangle from one's status within the hierarchy (Heffner, 1979). Who will be named as an author and who will be in the acknowledgements of a publication? Credit attribution criteria do not only rely on the nature of the contribution made, and numerous other factors come into play, namely the disciplinary context and sociodemographic variables which have been shown to affect one position within the hierarchy, such as gender, age and academic status (Merton, 1973; Cole and Cole, 1973; Zuckerman, 1977).

Heffner (1979) was one of the few to investigate scientific credit attribution using acknowledgements. Based on a questionnaire completed by 207 addressees from social and natural sciences, Heffner found that publication credit is not always accorded on the basis of the norm of universalism. More specifically, he showed that female PhDs were twice more likely than any other group (male and non-PhDs) to believe that they had been excluded from co-authorship when they felt their contribution warranted a place in the byline. Along those lines, Moore (1984) investigated the effect of authors' gender on the content of their acknowledgements and, more specifically, on the gender of those acknowledged. Using a sample of 300 male-authored and 70 female-authored psychology books, Moore found that

while men acknowledged mainly other men for their contribution and advice, women acknowledged the contributions of both genders. In another analysis, based on 684 psychology articles, Moore (1984) found a lower proportion of female acknowledgees, especially among articles from male authors. The author concludes that “there is a tendency on the part of each sex to seek out and acknowledge the professional advice of same-sex colleagues. For women, this most often means acknowledging individuals of both sexes (a practice relatively infrequent among male authors)” (p. 1029). More recently, Sugimoto and Cronin (2012) came to the same conclusion while analyzing the scholarly production of six important information scientists. They found that the authors included in their sample were more likely to acknowledge individuals of the same sex. Since those, few studies have focused on the issue of gender in credit attribution practices using acknowledgment data.

#### *Objective and Research Questions*

Focusing on the individuals named in acknowledgments, the objective of this work-in-progress is to better understand credit attribution practices of researchers and how gender may influence those. More specifically, we aim at answering the following research questions:

- What proportion of acknowledgees are women?
  - How does it vary as a function of the proportion of women in the byline of the acknowledging paper?
  - How does it vary as a function of the gender of the leading authors?
  - How does it vary by discipline?

#### **Material and methods**

##### *Data*

This study is based on all acknowledgements extracted from 2015 articles and reviews indexed in the Science Citation Index Expanded (SCI-E) and Social Sciences Citation Index (SSCI) from Clarivate Analytics' Web of Science (WoS). Access to the WoS data in a relational database format was provided by the *Observatoire des sciences et des technologies* (<http://www.ost.uqam.ca>). Acknowledgements data are collected and indexed by WoS only if they include funding information (Paul-Hus, Desrochers and Costas, 2016). These data are structured in three fields: the 'Funding Text' (FT), 'Funding Agency' (FO) and 'Grant Number' (FG). The dataset used in the present analysis was extracted from the FT field, which is the full text as it appears in the paper from which it is retrieved, and includes a total of 1,009,411 acknowledgements texts from as many papers. The dataset covers all disciplines included in the SCI-E and SSCI: Biology, Biomedical Research, Chemistry, Clinical Medicine, Earth and Space, Engineering and Technology, Health, Mathematics, Physics, Professional Fields, Psychology and Social Sciences. The discipline of a paper was assigned using the NSF field classification of journals (National Science Foundation, 2006); the NSF classification assigns only one discipline specialty to each journal, thus preventing potential multiple counting of papers.

##### *Analysis*

The extraction of individual names from acknowledgement texts was done using the Stanford Named Entity Recognizer (NER) (Finkel et al., 2005) module of the Natural Language ToolKit (NLTK) (Bird, 2009). To obtain the number of acknowledgee names per paper, the algorithm was applied on each string of acknowledgement text retrieved from the FT field and all named entities tagged as 'person' were selected.<sup>1</sup>

---

<sup>1</sup> This extraction procedure was used on a previous round of analysis of this dataset (Paul-Hus, Mongeon, Sainte-Marie & Larivière, 2017).

Several data cleaning procedures were then undertaken in order to eliminate non-human entities from the list of extracted names. First, incomplete names were removed from the list (entities containing only a first or last name, or only initials), retaining only entities composed of a complete name (i.e. full first and last name). To remove all remaining names that did not refer to individual persons such as grant, foundation, organization and institution names, manual cleaning was performed on the list. Examples of such names removed by manual cleaning include: Frederick Banting (grant), Marie Curie (grant and foundation), Boehringer Ingelheim (organization) and Instituto de Salud Carlos III (institution). Since acknowledgements often contain the name(s) of the author(s) signing the paper from which the acknowledgements are retrieved, a final step of cleaning was necessary. When the name(s) extracted from the acknowledgements of a paper X matched the name of one of the authors appearing in the byline of that paper X (using the first initial and the last name), this name was removed from the acknowledgees list for that specific paper, such as in the example below:

Paper X

Authors: J. Zhang, X. Feng and Y. Xu

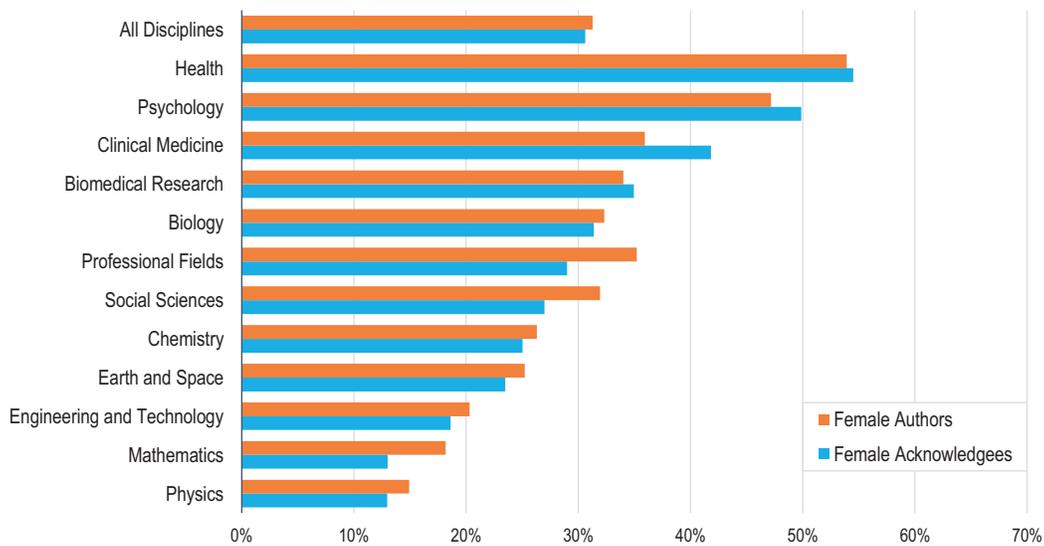
Acknowledgements text: “Jinsong Zhang, Xiao Feng, and Yong Xu contributed equally to this work [...]”

For the purpose of our analysis, we consider first and corresponding authors as lead authors of a paper since first authors are often associated with the highest proportion of tasks performed in a paper (Larivière et al., 2016) and corresponding authors—who are in charge of correspondence—are often associated to conception and supervision (Mattsson et al., 2011). If both the first and corresponding authors are women, the paper is considered female-led, if both are men, the paper is considered male-led and if first and corresponding author are of different genders, the paper is considered mixed.

The gender assignation of personal names (authors and acknowledgees) was done using the Wiki-Gendersort algorithm (Bérubé, Ghiasi and Larivière, in preparation). Using Wikipedia pages to get gender information, this algorithm increases reliability of gender assignation by examining the first names of the names covered by Wikipedia and counting the number of masculine and feminine pronouns in the introduction section of the first twenty pages. Gender is assigned to the first name when the same gender was attributed to 75% of Wikipedia pages. No gender is assigned when this threshold is not met. Using the Wiki-Gendersort algorithm, we were able to identify the gender of 75% of personal names in our dataset (authors and acknowledgees). The final dataset includes 878,250 acknowledgees mentioned in 291,167 papers.

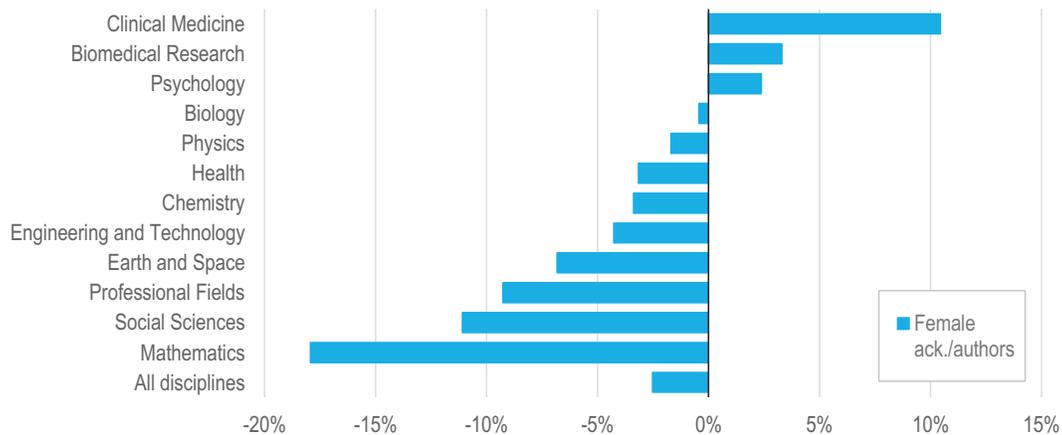
## Results

Figure 1 presents the global share of women among all authors and acknowledgees by discipline and shows that the well-known gender gap found in authorship (Larivière et al. 2013; West et al., 2013) is also present in the acknowledgements. Women represent less than 50% of authors and of acknowledgees in all disciplines, with the only exception of Health where women account for 54% of authors and 55% of acknowledgees. Moreover, Health, Psychology, Clinical Medicine and Biomedical Research—domains where women have been traditionally more present (Witz, 1992)—are the only disciplines where female acknowledgees represent a greater proportion than female authors. In all other disciplines, women are not only under-represented in authorships, they are even less present when looking at the acknowledgements.



**Figure 1. Proportion of female authors and acknowledgees, by discipline.**

To further investigate the representation of women, we analysed the proportion of female acknowledgees per paper in relation to the proportion of women among authors of a paper. If the share of female acknowledgees is higher than the share of women among authors, female acknowledgees are considered over-represented in relation to authorship, and inversely. Figure 2 demonstrates that in most disciplines, women are more under-represented in the acknowledgements as compared to their proportional presence in authorships. On average, when all disciplines are considered, women are underrepresented by 3% as compared to their share of authorships. Clinical Medicine, Biomedical Research and Psychology are the only disciplines where the trend is reverse and where women are overrepresented in the acknowledgements as compared to their proportion of authorships. Altogether, Figure 2 shows that, in most disciplines, women are even more under-represented in the acknowledgements than they are in authorships.



**Figure 2. Proportion of female acknowledgees as a function of the proportion of female authors, by discipline.**

Table 1 presents the proportion of acknowledgees who are women as a function of the gender of the leading authors. For all disciplines, the proportion of female acknowledgees is higher in female-led papers (women as first or corresponding authors) than in male-led papers or in the mixed group. The difference in the proportion of female acknowledgees between female and male-led papers ranges from 20.8% in the Professional Fields to 6.7% in Physics with a difference of 16.3% when all disciplines are taken together.

**Table 1. Proportion of acknowledgees who are female as a function of the gender of the leading authors, by discipline.**

Discipline	Female-led	Male-led	Mixed-led	Difference between female and male-led
Professional Fields	43.7%	22.9%	26.9%	20.8%
Health	61.3%	44.2%	52.3%	17.1%
Social Sciences	38.8%	22.0%	28.8%	16.7%
Clinical Medicine	52.1%	35.5%	43.7%	16.6%
Psychology	57.0%	41.4%	52.2%	15.6%
Chemistry	36.3%	21.8%	28.9%	14.5%
Mathematics	24.1%	10.5%	23.4%	13.7%
Engineering and Technology	28.1%	15.9%	26.5%	12.2%
Biomedical Research	43.1%	31.5%	37.9%	11.6%
Biology	37.3%	28.5%	35.3%	8.7%
Earth and Space	29.3%	21.6%	32.2%	7.7%
Physics	18.2%	11.5%	22.0%	6.7%
All Disciplines	42.0%	25.7%	36.9%	16.3%

### Discussion and conclusion

Our results demonstrate that gender disparities generally found in authorship extend to acknowledgements. In most disciplines, women are proportionally more under-represented in the acknowledgement than in the authors' list of a paper. Furthermore, as previously found by Moore (1984) and Sugimoto and Cronin (2012), our findings clearly confirm that women acknowledge proportionally more women than men. These results seem to indicate that gender plays a role in credit attribution practices. In a broader context, the under-representation of women in the acknowledgements of scientific papers adds up to the global gender disparities found in authorships, citations and self-citations (Ghiasi, Mongeon, Sugimoto and Larivière, 2018; Larivière et al., 2013; West et al., 2013).

To further our understanding of credit attribution practices, the next steps of this analysis will focus on the relationship between acknowledgements and academic status, looking more specifically at the number of papers published by acknowledgees and the number of citations received as a proxy for the reputation of a researcher.

### References

- Bérubé, N., Ghiasi, G., & Larivière V. (in preparation). Assigning gender to first names using Wikipedia.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly Media, Inc.

- Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information*, 14(6), 19-47.
- Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago, IL: University of Chicago Press.
- Finkel, J. R., Grenager, T., & Manning, C. (2005). Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics* (p. 363–370). Stroudsburg, PA, USA: Association for Computational Linguistics. <https://doi.org/10.3115/1219840.1219885>
- Ghiasi, G., Mongeon, P., Sugimoto, C.R., & Larivière, V. (2018). Gender homophily in citations. In *Proceedings of the 23rd International Conference on Science and Technology Indicators (p.1520-1525)*. Leiden, NL: Centre for Science and Technology Studies (CWTS). <http://hdl.handle.net/1887/65291>
- Heffner, A. G. (1979). Authorship Recognition of Subordinates in Collaborative Research. *Social Studies of Science*, 9(3), 377-384. <https://doi.org/10.1177/030631277900900305>
- Hoder-Salmon, M. (1978). Collecting Scholar's Wives. *Feminist Studies*, 4(3), 107-114. <https://doi.org/10.2307/3177543>
- Mattsson, P., Sundberg, C.J., and Laget, P. (2011). Is correspondence reflected in the author position? A bibliometric study of the relation between corresponding author and byline position. *Scientometrics*, 87, 99–105.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. Chicago and London: Chicago University Press.
- Moore, M. (1984). Sex and Acknowledgements: A Nonreactive Study. *Sex Roles*, 10 (11-12), 1021-1031.
- National Science Foundation. (2006). *Science and Engineering Indicators*. Chapter 5: Academic Research and Development. Data and Terminology. <http://www.nsf.gov/statistics/seind06/c5/c5s3.htm#sb1>
- Larivière, V., Ni, C., Gingras, Y., Cronin, B. & Sugimoto, C.R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504(7479), 211-213
- Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., and Sugimoto, C.R. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Sciences*, 46, 417–435.
- Paul-Hus, A., Mongeon, P., Sainte-Marie, M., & Larivière, V. (2017). The sum of it all: Revealing collaboration patterns by combining authorship and acknowledgements. *Journal of Informetrics*, 11(1), 80-87. <https://doi.org/10.1016/j.joi.2016.11.005>
- Paul-Hus, A., Desrochers, N., & Costas, R. (2016). Characterization, description, and considerations for the use of funding acknowledgement data in Web of Science. *Scientometrics*, 108(1), 167-182. <https://doi.org/10.1007/s11192-016-1953-y>
- Sugimoto, C. R., & Cronin, B. (2012). Biobibliometric profiling: An examination of multifaceted approaches to scholarship. *Journal of the American Society for Information Science and Technology*, 63(3), 450-468. <https://doi.org/10.1002/asi.21695>
- West, J.D., Jacquet, J., King, M.M., Correll, S.J., Bergstrom, C.T. (2013). The role of gender in scholarly authorship. *PLoS ONE* 8(7): e66212. <https://doi.org/10.1371/journal.pone.0066212>
- Witz, A. (1992). *Professions and Patriarchy*: Routledge.
- Zuckerman, H. A. (1977). *Scientific elite: Nobel laureates in the United States*. Piscataway, N.J.: Transaction Publishers.