

Gender and language use in scientific grant writing

Mackenzie Urquhart-Cronish^{ab*} and Sarah P. Otto^{ac}

^aBiodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada;

^bDepartment of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; ^cDepartment of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

*mackenzie.urq@gmail.com

Abstract

Women in science, technology, engineering, and math are not equally represented across tenure-track career stages, and this extends to grant funding, where women applicants often have lower success rates compared with men. While gender bias in reviewers has been documented, it is currently unknown whether written language in grant applications varies predictably with gender to elicit bias against women. Here we analyse the text of ~2000 public research summaries from the 2016 Natural Sciences and Engineering Research Council (NSERC) individual Discovery Grant (DG) program. We explore the relationship between language variables, inferred gender and career stage, and funding levels. We also analyse aggregated data from the 2012–2018 NSERC DG competitions to determine whether gender impacted the probability of receiving a grant for early-career researchers. We document a marginally significant gender difference in funding levels for successful grants, with women receiving \$1756 less than men, and a large and significant difference in rejection rates among early-career applicants (women: 40.4% rejection; men: 33.0% rejection rate). Language variables had little ability to predict gender or funding level using predictive modelling. Our results indicate that NSERC funding levels and success rates differ between men and women, but we find no evidence that gendered language use affected funding outcomes.

Key words: language, gender, career stage, funding, bias, NSERC

Introduction

Women in science face additional barriers to success that are rooted in historical biases (Wellenreuther and Otto 2016), which remain widespread despite over a decade of policies aimed at increasing female participation in science, technology, engineering, and math (STEM) fields (Larivière et al. 2013). While higher levels of gender parity have been achieved at earlier STEM career stages (e.g., graduate school, post-doctoral fellowships) (National Science Foundation and National Center for Science and Engineering Statistics 2018), there are still fewer women at senior levels (Larivière et al. 2013). This pattern has been described as the “glass obstacle course”, a metaphor for the unequal gendered processes at work in women’s careers in STEM fields (De Welde and Laursen 2011).

Gender diversity contributes to a diversity of perspectives, which can benefit scientific discovery by generating novel research questions and methods and facilitating wider application of research findings (Schiebinger et al. 2011–2018). The untapped potential of fully trained and credentialed female scientists limits advances in basic and applied research, and these limitations represent an



Citation: Urquhart-Cronish M and Otto SP. 2019. Gender and language use in scientific grant writing. FACETS 4: 442–458. doi:[10.1139/facets-2018-0039](https://doi.org/10.1139/facets-2018-0039)

Handling Editor: Imogen Coe

Received: September 25, 2018

Accepted: April 30, 2019

Published: September 2, 2019

Copyright: © 2019 Urquhart-Cronish and Otto. This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Published by: Canadian Science Publishing

important economic loss due to delayed improvements in technology, healthcare, and education (National Research Council 2010).

One obstacle that may face women scientists, particularly early in their careers, is a gender difference during grant review. For example, a meta-analysis conducted by Bornmann et al. (2007) reported that among grant applicants, men had 7% greater odds of being awarded funding than women. These gender differences in funding success could be related to various factors affecting the grant review process, including structural biases related to how academic scientific institutions function (e.g., there are more men in decision-making roles); gender biases (Kaatz et al. 2014), either explicit (e.g., conscious bias) or implicit (e.g., unconscious bias); and (or) differences in field, career, stage, or scientific productivity (e.g., Symonds et al. (2006) found that women had fewer papers early in their careers, but publication trajectories rose at a similar rate per year and citations per paper were higher for women). We set out to determine whether gender differences in scientific funding exist in Canada and whether language use in scientific grant proposals is associated with gender and funding level.

Written gendered language differences (e.g., written language characteristics associated with a particular sex or social gender) have been documented in nonscientific (Newman et al. 2008) and scientific contexts (Tse and Hyland 2008). For example, in reflective writing (e.g., narrative essays), female medical students used more words related to positive emotions than male students, and male medical students wrote longer documents compared with female students (Lin et al. 2016). It is currently unknown if differences in gendered language use in scientific grant writing exist, and no direct investigation of language use and gender for senior-level funding applications in STEM has been conducted. If a clear signature of gendered language use is detected and is related to funding differences between male and female researchers, steps can be taken to empower researchers with this knowledge and (or) to instruct reviewers to be careful about implicit bias being triggered by word choice. Conversely, if no differences are found then bias is unlikely to be triggered by linguistic associations with gender.

Here we investigate gender, career stage, and language use in publicly available summaries of individual Discovery Grant (DG) research proposals that were successfully funded in the 2016 awards competition by the Natural Sciences and Engineering Research Council (NSERC). NSERC funds research programs, not specific research projects, and DGs are the main source of general research funds for Canadian academics in the natural sciences and engineering. Principle investigators can only hold one DG at any point in time, and this program is important for sustaining research programs for Canadian scientists. Funding decisions and amounts are based on scores on: (i) excellence of the researcher, (ii) merit of the proposal, and (iii) training of highly qualified personnel and proposed budgets do not generally affect funding amounts, unless unjustified or underestimated (NSERC 2018). We use aggregated data supplied by NSERC from the 2012–2018 DG cycles to investigate the impact of gender on grant success and award levels for early-career researchers (ECRs, defined by NSERC as within three years of their first academic appointment). We then use all of the DG public summaries from 2016, as well as information about award values, to investigate the impact of language on grant funding.

Specifically, we use these data to address the following four questions:

1. Are gender and career stage associated with NSERC award value?
2. Do gender differences exist in language use within NSERC public summaries?
3. How does NSERC award value relate to language use, gender, and career stage?
4. What factors, including language use, predict career stage?

Materials and methods

The data used in this study are publicly available or are obtainable from NSERC in summary form, so no research ethics board approval was required (Article 2.2 of [CIHR, NSERC, and SSHRC 2014](#)).

Data and analysis of NSERC funding success and award levels

We requested funding data from NSERC, who provided aggregated data summarizing the total numbers of applicants and funded grants ([Table 1](#)). Because of an observed gender difference in funding success for ECRs in 2016, we asked for further data on ECR success rates from the 2012–2018 DG grant cycles, broken down by gender and by selection committee ([Table 2](#)). For this data, gender was self-reported by the grant applicants to NSERC (81.2% response rate in 2016; 80.3% response rate for ECRs in 2012–2018). Results that refer to the aggregated data in [Tables 1](#) and [2](#) use NSERC’s definition of ECR, with ECR status and gender self-reported to NSERC.

Data collection for linguistic data

Using R (version 3.3.3; [R Core Team 2017](#)) within the RStudio environment ([RStudio Team 2016](#)) and custom scripts, we collected all researcher data and NSERC public summaries ($n = 2094$) for the 2016 DG Competition from the NSERC database ([nserc-crsng.gc.ca/ase-oro/index_eng.asp](#), accessed 24 August 2017). We analyzed NSERC grants designated as “RGPIN” (research grant program to individuals) for the 12 main disciplines (i.e., “selection committee” as listed in [Table 2](#)).

Table 1. Number of applications and awards in the [NSERC \(2016\)](#) Discovery Grant cycle.

	All researchers	Early-career researchers
Number of applications		
Female	567 (17.8%)	106 (21.3%)
Male	2023 (63.4%)	274 (55.0%)
Not indicated	601 (18.8%)	118 (23.7%)
Total	3191	498
Number of awards		
Female	366 (17.3%)	74 (19.9%)
Male	1365 (64.5%)	208 (55.9%)
Not indicated	384 (18.2%)	90 (24.2%)
Total	2115	372
Award amount		
Female	\$32 342	\$26 366
Male	\$33 954	\$27 191
Not indicated	\$31 092	\$26 011
Total	\$33 155	\$26 741

Note: Includes Discovery and Subatomic Physics Individual and Team Grants, which were excluded from our analyses. Here, “Female”/“Male”/“Not indicated” or “Early-career” (within three years of securing an academic job) were self-reported to NSERC during the grant application stage, with exact numbers obtained directly from NSERC. NSERC, Natural Sciences and Engineering Research Council.

Table 2. NSERC selection committee codes and aggregated count data on early-career researcher applicants for Discovery Grant competitions in 2012–2018.

Selection committee	Awarded	Not awarded	Total
All early-career researchers			
Female	520	352	872
Male	1334	656	1990
Not indicated	492	210	702
1501: Genes, Cells, and Molecules			
Female	74	71	145
Male	224	130	354
Not indicated	44	17	61
1502: Biological Systems and Functions			
Female	121	118	239
Male	202	96	298
Not indicated	43	32	75
1503: Evolution and Ecology			
Female	40	35	75
Male	53	42	95
Not indicated	36	14	50
1504: Chemistry			
Female	20	15	35
Male	58	36	94
Not indicated	22	6	28
1505: Physics			
Female	27	10	37
Male	78	27	105
Not indicated	44	9	53
1506: Geosciences			
Female	41	22	63
Male	64	44	108
Not indicated	43	26	69
1507: Computer Science			
Female	28	12	40
Male	121	61	182
Not indicated	62	29	91
1508: Mathematics and Statistics			
Female	36	16	52
Male	103	44	147
Not indicated	55	18	73

(continued)

Table 2. (concluded)

Selection committee	Awarded	Not awarded	Total
1509: Civil, Industrial, and Systems Engineering			
Female	58	23	81
Male	154	51	205
Not indicated	53	21	74
1510: Electrical and Computer Engineering			
Female	22	6	28
Male	89	39	128
Not indicated	31	10	41
1511: Materials and Chemical Engineering			
Female	32	15	47
Male	75	37	112
Not indicated	20	9	29
1512: Mechanical Engineering			
Female	21	9	30
Male	113	49	162
Not indicated	39	19	58

Note: Gender (male, female, not indicated) was self-reported by applicants to NSERC, and “early-career researcher” was determined by NSERC. NSERC, Natural Sciences and Engineering Research Council.

Subatomic physics (individual and team grants) has different program designations and was not included in our analyses, reducing the number of awards to $n = 2021$ awardees.

We focussed our analysis on English public summaries ($n = 1959$), because the sample size for French public summaries was too small for an independent analysis ($n = 62$). Public summaries are limited in length by an online text box and English summaries were nearly 400 words in length (384.23 ± 1.6 words). In total, we analyzed 752 734 words across 1959 public summaries. Only the summaries and not the full grant proposals are publicly accessible; therefore, our linguistic analysis was limited to this portion of the grant application. Nevertheless, the public summaries form an important portion of the grant application, appearing on the first page of the proposal received by both the reviewers and the grant panel and serving as an abstract of the grant.

For each public summary, the gender of the author was inferred using the name association R package “gender” (Mullen 2015), which determines gender based on historical data sets of name use by gender, reporting the probability that a name is associated with a male (score of 1) or female (score of 0) individual. In cases where researcher gender was not strongly inferred (name not in the “gender” database or associated with gender probabilities between 0.2 and 0.8), we checked researcher websites and institutional news articles for identifying information (e.g., personal pronouns). We acknowledge that this method is subject to error and may not match each researcher’s self-identified gender (including nonbinary genders), except as captured by their current name choice, and report our results in an aggregate form only.

To determine career stage of each author, we requested information on year-of-first-grant for the successful 2016 DG cohort from the granting agency, but the request was declined citing logistical constraints. We thus inferred the career stage of each individual researcher by searching all years of data available in the NSERC online database (1991–2016; accessed 4 March 2018) for grants given to a researcher by the name and institution used in 2016. We calculated the number of years since researchers received their first listed NSERC DG as a discrete measure of career stage (range: 0–25 years). Our estimate of career stage was bounded by 25 years since the earliest date in the NSERC database was 1991, which may underestimate career stage for senior researchers. Researcher names and institutions were used jointly to limit misassignments of grants to different individuals with the same name, but we recognize that career stage is underestimated in cases where researchers change their name or institution, both within Canada and internationally. Our results using the individual data from the 2016 DG cycle (including all linguistic analyses) use inferred gender and inferred career stage (years since first NSERC DG at the same institution).

Generation of language variables

We used Language Inquiry and Word Count (LIWC) software (Pennebaker et al. 2015) to analyze the public summaries and generate language variables. LIWC software is widely used in the social sciences and has been parameterized across many different genres of writing, including scientific journal articles (Pennebaker et al. 2007), with the current version capturing, on average, over 86% of the words used in written text and speech (Pennebaker et al. 2015). Studies using LIWC have successfully detected gender differences in language in emails, narrative essays, and text excerpts from psychological studies (Newman et al. 2008; Cheng et al. 2011; Lin et al. 2016). LIWC is comprised of a large dictionary of words and compares inputted written text to its dictionary to generate scores for $n = 92$ language variables including word count, words per sentence, 86 traditional variables (e.g., content and style words), and four summary variables that are described below (Pennebaker et al. 2015).

Traditional LIWC variables include content and style words, and scores are expressed as a percentage of the total words used within the text provided. Content words generally include nouns, regular verbs, adjectives, and adverbs, whereas style words include pronouns, prepositions, articles, conjunctions, and auxiliary verbs. A broad distinction between the two word groups is that content words reflect what is being said, whereas style words reflect how people are communicating (Tausczik and Pennebaker 2010). The four summary variables included are Analytic thinking, Clout, Authenticity, and Emotional tone. Analytic thinking refers to the degree of formal, logical, or hierarchical thinking patterns in text. Clout scores writing that is authoritative, confident, and exhibits leadership. Authenticity refers to writing that is personal and honest. Finally, Emotional tone is scored such that higher numbers are more positive and lower numbers are more negative. These four summary variables are research-based composites that have been converted to 100-point scales, where 0 = very low along the dimension and 100 = very high (Pennebaker et al. 2015). Before data analysis, language variables ($n = 92$) were tested and removed if they exhibited near-zero variance (Kuhn 2017), leaving $n = 74$ language variables used in reported analyses.

Statistical analysis of award data

To investigate the magnitude of the gendered funding gap in the 2016 DG cohort we used a linear mixed-effects model relating award value to inferred gender and career stage. Gender and career stage were treated as fixed effects, whereas selection committee (i.e., discipline) was treated as a random effect to control for discipline-specific differences in average award value. Throughout, we used the Kenward–Roger adjustment for the degrees of freedom in the mixed-effects models.

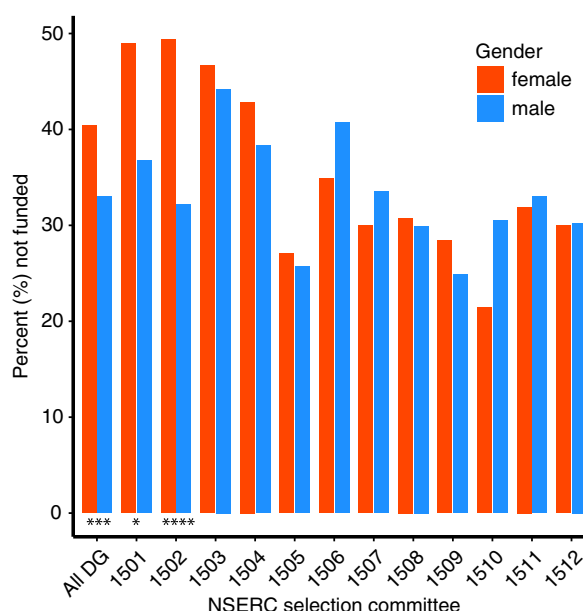


Fig. 1. The percentage of grants that remained unfunded was significantly lower from early-career male applicants (see [Table 1](#) for selection committees and count data). Asterisks indicate a significant difference in the proportion unfunded between male and female awardees (χ^2 test: *, $p < 0.05$; ***, $p < 0.001$; ****, $p < 0.0001$). Natural Sciences and Engineering Research Council's (NSERC) definition of early-career researcher has changed over this time period but has consistently required applicants to be within the first few years of their appointments. Some applicants may have applied and been evaluated in multiple years (these cases could not be distinguished in the summary data provided by NSERC). Gender was self-identified to NSERC; 19.7% of the applicants chose not to specify their gender (data not analyzed).

To determine if funded versus not-funded outcomes depended on gender, we performed Pearson's χ^2 tests on the total data and for each selection committee separately, based on the aggregated data from 2012 to 2016 ([Fig. 1](#)). We also performed a Mantel-Haenszel χ^2 test, which accounts for variation among selection committees and tests the null hypothesis of a common odds ratio of one (female and male researchers equally likely to be funded). Odds ratios were calculated using the R package “samplesizeCMH” in R version 3.4.3.

Statistical analysis of linguistic data

We used a principal component analysis (PCA) based on covariances to visualize the linguistic data and used a permutational multivariate analysis of variance using distance matrices (PERMANOVA) from the R package “vegan” ([Oksanen et al. 2017](#)) to examine whether gender explains differences in multivariate LIWC variables. To determine how many axes were meaningful, we compared the observed eigenvalues to a broken-stick random expectation; axes in which the observed eigenvalue was greater than randomly generated expected eigenvalues were considered meaningful. The broken-stick model has been shown to yield more consistent results compared to other stopping methods ([Jackson 1993](#)). Our comparison identified the first seven principal component (PC) axes as more informative than expected, explaining 40% of the variance in the data.

We used language variables with machine-learning techniques and random-forest predictors—a model averaging approach—to train and test classification and regression models using the R package “caret” ([Kuhn 2017](#)). The data were randomly subdivided into a training data set (75%) and a testing data set (25%). To address our imbalanced gender data and avoid classification models that always predict the most common class (men), we up-sampled women in our training data to an equivalent sample size of men (i.e., adding women by randomly sampling them from the training data set with replacement). The trained model was then applied to the testing data set, with its original unbalanced gender composition. We used one classification model with gender as a binary response variable and two regression models with career stage and award value as continuous response variables. We conducted mixed-effects models on variables that explained the most variance in

our random-forest models, accounting for selection committee and gender (as appropriate) and controlling for multiple comparisons using a Bonferroni correction.

In addition to the presented results, we conducted analyses on data sets with reduced language variables to assess the robustness of our results, either using only the first seven PC axes or restricting the language variables to those exhibiting a standard deviation >1 (see [Supplementary Material 1](#)). We found no major differences in model performance compared with the broad language data set analysed here.

Results

1) Are gender and career stage associated with NSERC award value?

We used the aggregated data from NSERC for the 2016 DG cycle to determine differences in success rate by gender ([Table 1](#)). Overall, the success rates for applicants who self-identified as male and female were 64.6% and 67.5%, respectively, amounting to a higher success rate for men relative to women by a factor of 1.045 ([Table 1](#)). The average grant size was \$33 155 ([NSERC 2016](#)), with men receiving a factor of 1.050 more funding relative to women ([Table 1](#)).

Analysing the individual data from the 2016 DG cycle, we inferred that 21.4% of awardees were women and 78.6% men, consistent with the aggregated data from 2016 ([Table 1](#); 21.1% women and 78.9% men among those who self-reported gender) and from previous years (table 4.1 in [NSERC 2017](#)). The proportion of female awardees within each selection committee ranged from ~10% in the physical sciences, engineering, and mathematics to ~30% in life and Earth sciences ([Fig. S1](#)). We inferred that 36% of researchers received their first DG award in the 2016 competition, which is likely an overestimate given name and institutional changes. Indeed, this estimate is substantially higher than the 15.0% of awardees reported to be ECRs ([Table 1](#)), although a similar fraction of awardees were “Established Researchers Not Holding a Grant” (table 2 in [NSERC 2016](#)), who would also contribute to the number of awardees receiving a DG for the first time. Nevertheless, we caution that our measure of career stage, as inferred here, only reflects estimated time since first receiving an NSERC DG and is biased downwards by name and institutional changes.

Analysing the individual data for 2016 DG awardees, the average annual grant was \$34 375 (± 623 SE) for women and \$36 264 (± 384 SE) for men. Only career stage significantly affected differences in amount of funding awarded ($F_{[1,1946.2]} = 66.50, p < 0.0001$), with a marginally significant effect of gender ($F_{[1,1945.3]} = 2.89, p = 0.089$), and no significant interaction between the two variables ($F_{[1,1944.6]} = 0.02, p = 0.902$) ([Table 3](#)). Keeping in mind the caveat of our measure of career stage,

Table 3. Summary statistics from linear mixed-effects models investigating the effects of gender and career stage on award value.

Variable	All researchers (n = 1959)				Awarded first Natural Sciences and Engineering Research Council Discovery Grant in 2016 (n = 710)			
	Estimate [95% CI]	SE	df	p	Estimate [95% CI]	SE	df	p
Gender	1756.2 [−267.2, 3779.5]	1032.3	1	0.089	1620.2 [−42.5, 3282.8]	848.3	1	0.057
Career stage	386.4 [221.8, 551.0]	84.0	1	<0.0001	—	—	—	—
Gender × career stage	−11.40 [−192.7, 169.9]	93.48	1	0.902	—	—	—	—

Note: Estimates for gender are reported as the funding level for male awardees minus that to female awardees. Estimates for career stage are reported as the increase in funding level with each additional year since an applicant’s first award (see Methods for caveats).

the main effect of career stage amounted to an average \$386 (\pm 83.97 SE) increase in award value for every one-year increase in years since first DG award. Although only marginally significant, the gender difference in award amount, accounting for career stage and selection committee, was \$1756 (\pm 1032 SE).

Obtaining funding as an assistant professor is an important step in establishing an academic scientific career. Among the inferred first-time recipients, a strong male bias was also found in the 2016 cohort; 77% of the first-time recipients were men and 23% women. Inferred first-time recipients who were men also received a higher average award of \$31 965 (\pm 469 SE), compared to \$30 257 (\pm 587 SE) for women. The gender bias in funding level for researchers inferred to have received their first NSERC DG in 2016 was again only marginally significant ($F_{[1,699.3]} = 3.65$, $df = 1$, $p = 0.057$) (Table 3). Specifically, the gender difference in award amount among inferred first-time recipients, controlling for selection committee, was \$1620 (\pm 848 SE), a roughly 5% higher award amount for male applicants relative to female applicants.

An even larger discrepancy exists when considering gender differences in the proportion of ECRs who were funded versus unfunded. Figure 12 in NSERC's (2016) annual summary suggested higher rejection rate for female ECRs relative to male ECRs. To investigate further, we requested data on summary outcomes for early-career applicants by gender and by selection committee from the 2012–2018 DG grant cycles (Table 2). These data were received in aggregated form and included ECR status and gender as self-reported to NSERC. Applications from female ECRs were rejected 40.4% of the time (352/872) compared to only 33.0% time for male ECRs (33.0%; 656/1990), a significant result with females being rejected 1.225 times more often ($p = 0.00016$ Pearson's χ^2 test), which remains significant when accounting for selection committee ($p = 0.0033$ Mantel–Haenszel χ^2 test). Figure 1 indicates that the gender disparity is particularly high in two of the selection committees (1501: Genes, Cells, and Molecules; 1502: Biological Systems and Functions).

2) Do gender differences exist in language use within NSERC public summaries?

We first performed a PCA on the linguistic data to focus on the main axes of variation among the multiple language variables measured by LIWC ($n = 74$). The variables with the largest loadings for PC1 were total pronouns, function words, dictionary words, the time orientation category “present focus”, and regular verbs all having negative loadings. The main loading variables for PC2 were Clout, first person plural, affiliation, and social words having positive loadings and affect words having a negative loading. Because of the large sample size of our data set, we observed a significant effect of gender on the first two PC axes ($F_{[1,1957]} = 4.86$, $p = 0.001$), with women tending to score higher on PC2 (Fig. 2). Nevertheless, gender explained little of the variability in language variables ($R^2 = 0.002$).

To investigate whether variation in language use could accurately predict author gender, we used a random-forest classification model using language variables to predict gender of award recipients, including selection committee as a covariate. Predictive model accuracy was lower than the no-information rate (NIR) (accuracy = 0.77, NIR = 0.79), and the random forest classification had a poor ability to correctly classify women (sensitivity = 0.06), despite up-sampling women to equal the sample size of men during model training, but a good ability to correctly classify men (specificity = 0.97). Precision, or a measure of classifier exactness, was 0.30, indicating a moderate false-positive rate. Cohen's kappa, a measure of classification accuracy normalized by the imbalance of the gender classes in the data, was 0.03, indicating a weak predictive ability. We confirmed our main results on predictive ability using another classification method (linear discriminant analysis also performed poorly: Cohen's kappa = 0.11).

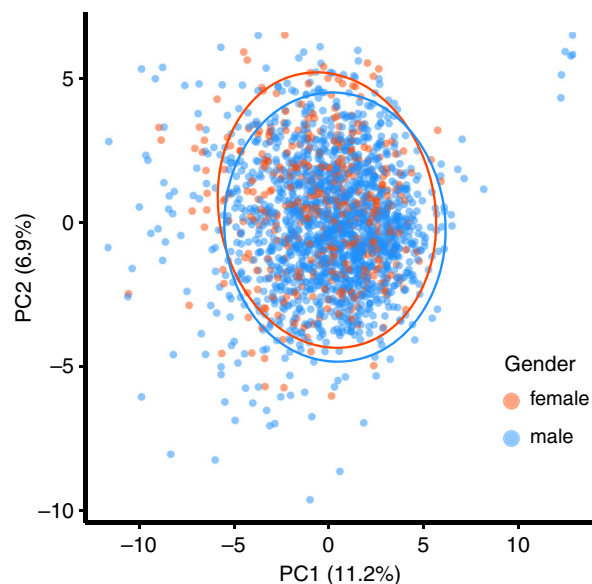


Fig. 2. Principal component (PC) ordination plot of the first two axes with ellipses around 95% confidence levels. PC1 explains 11.2% of the variation in language variables, while PC2 explains 6.9%. The next five PC axes explained between 5.5% and 3.5% of the variation in language variables.

Among the linguistic variables, “conjunctions” was the most important variable in our random-forest model with the highest predictive ability, explaining $\sim 15\%$ of the variation ascribed to gender. Considering the variable on its own, women had a significantly higher mean conjunction score than men in a mixed-effect model controlling for selection committee (women = 6.42, men = 6.00; $F_{[1,1955.2]} = 19.72$, $p < 0.0001$, which is less than the Bonferroni corrected $\alpha = 0.00068$ having considered 74 language variables). Conjunctions (e.g., “and”, “also”, “but”, “though”, “whereas”) are style words that reflect how people are communicating. Funding amount did not vary with conjunction score, controlling for gender and selection committee in a mixed-effects model ($F_{[1,1949.8]} = 0.267$, $df = 1$, $p = 0.605$).

3) How does NSERC award value relate to language use, gender, and career stage?

We were also interested in testing whether language variables, gender, and career stage were predictors of award value. Although language use did not predictably vary by gender, we investigated whether certain writing styles or word use accurately predicted funding level, when accounting for career stage. To account for variation in award value among disciplines, we first scaled award value by the average award within each selection committee. We then used a random-forest regression model to assess the effects of language variables, career stage, and gender on scaled award value. A random-forest regression model using the testing data set had the same low R^2 of 0.03 as the training data set, indicating a poor predictive ability, with a marginally higher root-mean-square error (RSME; testing: 14 495; training: 12 970) and mean absolute error (MAE; testing: 9920; training: 9374). Career stage was the most important predictor, explaining $\sim 9\%$ of the variation in scaled award value. Of the linguistic variables, “total pronouns” was the most important, explaining $\sim 6\%$ of the variation in scaled award value. Considering this linguistic variable on its own, “total pronouns” was not a significant predictor of award value, controlling for selection committee and gender, once adjusting for multiple comparisons ($F_{[1,1946.2]} = 8.70$, $p = 0.003$, which is much greater than the Bonferroni corrected $\alpha = 0.00068$ having considered 74 language variables). Pronouns include personal pronouns (e.g., “I”, “we”, “us”, “you”, “she”, “him”, “her”, “they”) and impersonal pronouns (e.g., “it”, “its”, “those”).

4) What factors, including language use, predict career stage?

Finally, we tested whether language variables could be used to predict career stage, for example, whether writing style, word use, or tone predictably changed with continued grant writing experience, treating gender and selection committee as covariates. A random-forest regression model using the testing data set had a slightly higher, but still low, R^2 (testing: 0.07; training: 0.05), with similar RSME (testing: 0.96; training: 0.98) and MAE (testing: 0.83; training: 0.84). The “Genes, Cells, and Molecules” selection committee emerged as the most important predictor variable, explaining ~6% of the variation in career stage of the 2016 awardees. According to our measure of career stage, more awardees had their first recorded DG award in 2016 in this selection committee (47%) compared with the total across all selection committees (36%). Among the linguistic variables, “tentative” was the most important predictor of career stage, explaining ~5% of the variation. Considering the language variable “tentative” on its own, tentative scores increased with career stage, accounting for selection committee and gender as random effects, but this was only marginally significant after correcting for multiple comparisons ($F_{[1,1921.1]} = 11.09$, $p < 0.0009$, which is slightly greater than the Bonferroni corrected $\alpha = 0.00068$ having considered 74 language variables). Note that if we further remove linguistic variables that exhibit little variance using a higher threshold (standard deviation ≤ 1), the rise in tentative score with career stage remains significant even after Bonferroni correction (see [Supplementary Material 1](#)). The tentative score measures use of words such as “maybe”, “perhaps”, or “guess”.

Discussion

Representation of women continues to lag in many scientific disciplines, with women particularly underrepresented among senior academic scientists ([Council of Canadian Academies 2012](#)). A report by the [Council of Canadian Academies \(2012\)](#) concluded that “... time alone will probably not be enough to balance the proportion of women and men at the highest levels of academia” (p. 53). The challenges that stymie career progress for women in STEM are multifaceted ([Shen 2013](#)) and include biases that diminish evaluations of women’s scientific work. We sought to determine whether grant funding differed by gender in Canada and whether language choice in scientific writing was correlated with gender, hypothesizing that men and women may differ in the tone of their word use (e.g., using words that are less authoritative in tone). If so, then biases could be triggered by reading scientific texts, even if the gender of the writer is not known. If such gendered word use could be detected, then scientists and reviewers might gain from the knowledge of how texts differ by gender and how they might be read. These questions are timely, as reflected by NSERC’s current Equity, Diversity, and Inclusivity initiatives to address issues of equal representation in STEM and increase the social relevance and impact of research ([Holmes and NSERC 2018](#)). Below we summarize our main results in light of the questions posed in the introduction.

Gender and career stage are associated with NSERC award value

Our results show the effect of gender on award value in 2016 was marginally significant once we accounted for career stage and selection committee, with women at the same career stage awarded ~5% smaller research grants than men (\$1756 less, [Table 3](#), based on inferred gender). Career stage had a significant effect on funding amount (see Results), with researchers awarded an average of \$386 more for every additional year since they first held an award (see also NSERC report 2016). Previous studies have also reported gender differences in STEM grant success rates ([Bornmann et al. 2007](#)) and funding amounts ([Head et al. 2013](#)).

We also found the success rate of the ECRs was significantly higher for men than women ([Fig. 1](#)), using NSERC’s definition of ECRs as within three years of their first academic appointment and based on the aggregated data supplied by NSERC (personal communication, 2018). Indeed, ECRs were significantly more likely to have their grant rejected if they were a woman (40.4% rejection rate) than

if they were a man (33.0%) over the 2012–2018 award cycles ([Table 2](#); [Fig. 1](#)). Given evidence that grant rejection at this early stage can have a substantial negative impact on future participation in funding competitions ([Bol et al. 2018](#)), the detected inequality in success rates poses a potentially important obstacle in women’s academic careers.

Determining the extent to which bias plays a role in the evaluation of research excellence at this early stage and the impact of grant rejection on reapplication rates for women should be a major research priority. We thus recommend that NSERC investigate the causes behind the large discrepancy in funding success between early-career men and women and institute policies that correct for any biases that may exist in the assessment of grants. Initiatives that prioritize efforts at early-career stages—especially those in underrepresented groups, including women—are particularly valuable, assisting the next generation of academics in getting their start.

Gender differences in language use were not detected within NSERC public summaries

Contrary to our expectation that scientific writing might be gendered, we found that public summaries of awarded grants did not differ substantially in language use between male and female academics, as measured using the LIWC program. While a small and significant difference could be detected between male and female authors ([Fig. 2](#)), there was little to no ability to predict gender based on language use, according to either a random-forest classifier or a linear discriminant analysis. It has been previously demonstrated that as writing becomes more technical, gendered language differences diminish ([Francis et al. 2001](#); [Yavari and Kashani 2013](#)). Our results similarly indicate little difference between male and female awardees in linguistic word use in formal scientific writing, at least as captured within NSERC public summaries ([Fig. 2](#)). This suggests that word choice in successfully funded grant applications is unlikely to trigger gender biases.

The most distinguishing variable was “conjunction”, which explained about 15% of the variation ascribed to gender. One possible explanation is a difference in narrative style between men and women. Conjunctions function to join multiple thoughts together and are important for creating a coherent narrative in writing ([Graesser et al. 2004](#)). Scientists are increasingly incorporating elements of narrative style in their writing, which has been associated with articles published in higher-impact journals ([Hillier et al. 2016](#)), a factor that may be related to increased funding success according to NSERC’s evaluation criteria ([NSERC 2018](#)). We found a significant difference in conjunction scores between men and women, with women scoring higher on average, demonstrating an aspect of language where gender differences exist.

NSERC award value or career stage were poorly predicted by language variables

We then examined whether award value could be predicted by differences in language use. For example, if certain styles of writing are associated with publications in higher-impact journals ([Hillier et al. 2016](#)), those styles also may be favoured by grant proposal reviewers. However, our random-forest analysis indicated that language use had little power to predict award value, with career stage being the most important predictor variable. This finding is consistent with the results of our linear mixed-effects model of the 2016 DG data, where we found a significant and positive relationship between time since first NSERC DG and award value, which may reflect increasing experience gained over the course of a career.

The language variable that explained the most variation in scaled award value in our model was “total pronouns”, including all personal and impersonal pronouns. First- and second-person pronouns are

associated with the active voice. As mentioned, narrative writing styles using the active voice have been associated with publications in higher-impact journals (Hillier et al. 2016), which could affect grant evaluation and subsequent award value. While award level rose with total pronoun score, the relationship was nonsignificant following Bonferroni correction.

Finally, we sought to investigate whether language use differed among researchers at different career stages. Career stage was poorly predicted by language variables. Instead, one of the covariates—the “Genes, Cells, and Molecules” selection committee—was the most important predictor variable, with more awardees inferred to be receiving a grant for the first time in 2016 (47% versus 36%, averaged across all selection committees). The language variable that explained the most variation in model performance was “tentative”, although the result was only marginally significant after Bonferroni correction. Tentative language tended to increase with career stage, which may reflect more cautious language use with respect to the proposed research and (or) an increased attention to gaps in the state of the field.

Strengths and limitations of our study

Although our predictive modelling framework found that linguistic variables were poor predictors of gender or grant value, we have no reason to believe our results are an artifact of the model training and tuning process. We up-sampled women in our classification training model to avoid biases in the model training process and performed multiple rounds of cross-validation without resampling to reduce variability and optimize model performance. We also have no reason to believe word count of analyzed public summaries ($n = 1959$) was too low (average of 384 words) to detect gender differences in writing, because machine-learning techniques have been able to detect gender differences in writing using emails with a mean word count of 116 using a training sample size of $n = 1000$ (Cheng et al. 2011).

Nevertheless, gendered language differences may still exist within STEM fields in other sections of written grant proposals, funded versus unfunded grants proposals, informal writing (e.g., email), at earlier career stages (e.g., graduate school or post-doctoral fellowships), or when evaluating aspects other than word use such as sentence structure and content. Thus, language use may still affect funding success beyond the scope of our analyses.

In particular, the public summaries that we analysed may have a more formulaic style, reducing gender differences in language use. A linguistic analysis of full grant proposals or the section “Most Important Contributions”, which describes past research accomplishments, might reveal more differences than are apparent in the public summaries (e.g., in personal pronoun use, emotional tone, clout), although neither are publicly available. Additionally, as unfunded grants are not included on NSERC’s public database, we could not analyze whether word use differed by gender and funding outcome (funded or unfunded) and could not determine the impact of a grant rejection on the proportion of men and women who re-apply. Future studies incorporating different types of writing, different academic career stages, and different success outcomes (e.g., funded vs. unfunded) would provide a more complete picture of the interaction between gender, career stage, and language use on research funding.

Conclusions

The current underrepresentation of women in senior scientific positions will not be solved without proactive policies (Holman et al. 2018; Grogan 2019). Pursuing potential factors driving biases (e.g., explicit, implicit, structural) that diminish evaluations of women’s scientific work is necessary to achieve equity. We identified a significant gender difference in success rates among ECRs using aggregated data provided directly by NSERC, with women remaining unfunded by a factor of 1.225

more than men (40.4% vs. 33.0% rejection rate, respectively). This early-career gender disparity is a potentially important obstacle impeding the research progress of female scientists and should be investigated.

We also examined whether language choice in scientific writing varied in predictable ways according to gender, hypothesizing that male and female academics may differ in aspects of their word use. Our predictive modelling framework found little evidence of distinct writing styles, tones, or word use separating male and female academics in the 2016 funded DG cohort. Our classification model did identify a minor difference in conjunction use as language variable that distinguished men and women's writing, with women having a slightly higher conjunction score on average (6.42 vs. 6.00).

Our study is, however, correlative, and we cannot determine the causes of the detected gender differences in award success and award amount. For example, we cannot distinguish implicit or explicit biases during the review process from differences in measures of scientific productivity. Further research that identifies and draws direct links between funding barriers and funding outcomes will strengthen equity, diversity, and inclusivity initiatives, thereby reducing obstacles facing the progress of underrepresented groups as they proceed through tenure-track career stages.

Acknowledgements

We would like to thank J.H. Myers for helpful discussion. R.M. Germain, L.M. Guzman, K.A. Thompson, S. Watt, and two anonymous reviewers provided helpful comments on an earlier draft of this manuscript. We thank C. Whillans and S. Overington at NSERC for providing the data on aggregated funded and unfunded grants for early-career researchers ([Table 2](#)); the conclusions emerging from this work are exclusively those of the authors and do not necessarily reflect the opinions of NSERC. This work was supported by an NSERC Discovery Grant to SPO (RGPIN-2016-03711).

Author contributions

MU-C and SPO conceived and designed the study. MU-C and SPO performed the experiments/collected the data. MU-C and SPO analyzed and interpreted the data. SPO contributed resources. MU-C and SPO drafted or revised the manuscript.

Competing interests

The authors have declared that no competing interests exist.

Data availability statement

All relevant data are within the paper, the Supplementary Material, and Dryad digital repository provisional doi:[10.5061/dryad.6mg5g10](https://doi.org/10.5061/dryad.6mg5g10).

Supplementary material

The following Supplementary Material is available with the article through the journal website at doi:[10.1139/facets-2018-0039](https://doi.org/10.1139/facets-2018-0039).

Supplementary Material 1

References

Bol T, de Vaan M, and van de Rijdt A. 2018. The Matthew effect in science funding. *Proceedings of the National Academy of Sciences of the USA*, 115: 4887–4890. PMID: [29686094](#) DOI: [10.1073/pnas.1719557115](#)

Bornmann L, Mutz R, and Daniel HD. 2007. Gender differences in grant peer review: a meta-analysis. *Journal of Informetrics*, 1: 226–238. DOI: [10.1016/j.joi.2007.03.001](https://doi.org/10.1016/j.joi.2007.03.001)

Canadian Institutes of Health Research (CIHR), Natural Sciences and Engineering Research Council of Canada (NSERC), and Social Sciences and Humanities Research Council of Canada (SSHRC). 2014. Tri-council policy statement: ethical conduct for research involving humans. Secretariat on Responsible Conduct of Research, Ottawa, Ontario.

Cheng N, Chandramouli R, and Subbalakshmi KP. 2011. Author gender identification from text. *Digital Investigation*, 8: 78–88. DOI: [10.1016/j.diin.2011.04.002](https://doi.org/10.1016/j.diin.2011.04.002)

Council of Canadian Academies. 2012. Strengthening Canada's research capacity: the gender dimension. Council of Canadian Academies, Ottawa, Ontario.

De Welde K, and Laursen SL. 2011. The glass obstacle course: informal and formal barriers for women Ph.D. students in STEM fields. *International Journal of Gender, Science and Technology*, 3: 571–595.

Francis B, Robson J, and Read B. 2001. An analysis of undergraduate writing styles in the context of gender and achievement. *Studies in Higher Education*, 26: 313–326. DOI: [10.1080/03075070120076282](https://doi.org/10.1080/03075070120076282)

Graesser AC, McNamara DS, Louwerse MM, and Cai Z. 2004. Coh-Metrix: analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, 36: 193–202. PMID: [15354684](https://pubmed.ncbi.nlm.nih.gov/15354684/) DOI: [10.3758/BF03195564](https://doi.org/10.3758/BF03195564)

Grogan KE. 2019. How the entire scientific community can confront gender bias in the workplace. *Nature Ecology & Evolution*, 3: 3–6. PMID: [30478306](https://pubmed.ncbi.nlm.nih.gov/30478306/) DOI: [10.1038/s41559-018-0747-4](https://doi.org/10.1038/s41559-018-0747-4)

Head MG, Fitchett JR, Cooke MK, Wurie FB, and Atun R. 2013. Differences in research funding for women scientists: a systematic comparison of UK investments in global infectious disease research during 1997–2010. *BMJ Open*, 3: e003362. PMID: [24327360](https://pubmed.ncbi.nlm.nih.gov/24327360/) DOI: [10.1136/bmjopen-2013-003362](https://doi.org/10.1136/bmjopen-2013-003362)

Hillier A, Kelly RP, and Klinger T. 2016. Narrative style influences citation frequency in climate change science. *PLoS ONE*, 11: e0167983. PMID: [27978538](https://pubmed.ncbi.nlm.nih.gov/27978538/) DOI: [10.1371/journal.pone.0167983](https://doi.org/10.1371/journal.pone.0167983)

Holman L, Stuart-Fox D, and Hauser CE. 2018. The gender gap in science: how long until women are equally represented? *PLoS Biology*, 16: e2004956. PMID: [29672508](https://pubmed.ncbi.nlm.nih.gov/29672508/) DOI: [10.1371/journal.pbio.2004956](https://doi.org/10.1371/journal.pbio.2004956)

Holmes A, and NSERC. 2018. Toward a new normal. Equity, diversity and inclusion as integral to research and innovation excellence: models for success. Outcomes from Gender Summit 11 North America, Montreal, Quebec, 6–8 November 2017. NSERC, Ottawa, Ontario. 20 p. [online]: Available from gender-summit.com/gsl1-about.

Jackson DA. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology*, 74: 2204–2214. DOI: [10.2307/1939574](https://doi.org/10.2307/1939574)

Kaatz A, Gutierrez B, and Carnes M. 2014. Threats to objectivity in peer review: the case of gender. *Trends in Pharmacological Sciences*, 35: 371–373. PMID: [25086743](https://pubmed.ncbi.nlm.nih.gov/25086743/) DOI: [10.1016/j.tips.2014.06.005](https://doi.org/10.1016/j.tips.2014.06.005)

Kuhn M. 2017. caret: classification and regression training. R package version 6.0-77. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew

Ziem, Luca Scrucca, Yuan Tang, Can Candan, and Tyler Hunt [online]: Available from CRAN.R-project.org/package=caret.

Larivière V, Ni C, Gingras Y, Cronin B, and Sugimoto CR. 2013. Bibliometrics: global gender disparities in science. *Nature*, 504: 211–213. PMID: [24350369](#) DOI: [10.1038/504211a](#)

Lin C-W, Lin M-J, Wen C-C, and Chu S-Y. 2016. A word-count approach to analyze linguistic patterns in the reflective writings of medical students. *Medical Education Online*, 21: 29522. PMID: [26838331](#) DOI: [10.3402/meo.v21.29522](#)

Mullen L. 2015. gender: predict gender from names using historical data. R package version 0.5.1 [online]: Available from cran.r-project.org/web/packages/gender/index.html.

National Research Council. 2010. Gender differences at critical transitions in the careers of science, engineering, and mathematics faculty. The National Academies Press, Washington, D.C.

National Science Foundation and National Center for Science and Engineering Statistics. 2018. Doctorate recipients from U.S. universities: 2016. Special Report NSF 18-304. Alexandria, Virginia [online]: Available from nsf.gov/statistics/2018/nsf18304/.

Newman ML, Groom CJ, Handelman LD, and Pennebaker JW. 2008. Gender differences in language use: an analysis of 14,000 text samples. *Discourse Processes*, 45: 211–236. DOI: [10.1080/01638530802073712](#)

NSERC. 2016. 2016 competition statistics discovery grants (DG) and research tools and instruments (RTI) programs. Natural Sciences and Engineering Research Council of Canada, Ottawa, Ontario. 22 p.

NSERC. 2017. Women in science and engineering in Canada. Natural Sciences and Engineering Research Council of Canada, Ottawa, Ontario. 75 p.

NSERC. 2018. Discovery grants peer review manual. Natural Sciences and Engineering Research Council of Canada, Ottawa, Ontario. 45 p.

Oksanen J, Guillaume Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, et al. 2017. vegan: community ecology package. R package version 2.4-5 [online]: Available from CRAN.R-project.org/package=vegan.

Pennebaker JW, Chung CK, Ireland M, Gonzales A, and Booth RJ. 2007. The development and psychometric properties of LIWC2007. University of Texas at Austin, Austin, Texas.

Pennebaker JW, Boyd RL, Jordan K, and Blackburn K. 2015. The development and psychometric properties of LIWC2015. University of Texas at Austin, Austin, Texas.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [online]: Available from R-project.org.

RStudio Team. 2016. RStudio: integrated development for R. RStudio, Inc., Boston, Massachusetts [online]: Available from rstudio.com/.

Schiebinger L, Klinge I, Sánchez de Madariaga I, Paik HY, Schraudner M, and Stefanick M (*Editors*). 2011–2018. Gendered innovations in science, health & medicine, engineering, and environment [online]: Available from genderedinnovations.stanford.edu.

Shen H. 2013. Inequality quantified: mind the gender gap. *Nature*, 495: 22–24. PMID: [23467149](#) DOI: [10.1038/495022a](#)

Symonds MRE, Gemmell NJ, Braisher TL, Gorringer KL, and Elgar MA. 2006. Gender differences in publication output: towards an unbiased metric of research performance. *PLoS ONE*, 1: e127. PMID: [17205131](#) DOI: [10.1371/journal.pone.0000127](#)

Tausczik YR, and Pennebaker JW. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29: 24–54. DOI: [10.1177/0261927X09351676](#)

Tse P, and Hyland K. 2008. ‘Robot Kung fu’: gender and professional identity in biology and philosophy reviews. *Journal of Pragmatics*, 40: 1232–1248. DOI: [10.1016/j.pragma.2007.02.002](#)

Wellenreuther M, and Otto S. 2016. Women in evolution—highlighting the changing face of evolutionary biology. *Evolutionary Applications*, 9: 3–16. PMID: [27087836](#) DOI: [10.1111/eva.12343](#)

Yavari M, and Kashani AF. 2013. Gender-based study of metadiscourse in research articles’ rhetorical sections. *International Journal of Applied Linguistics & English Literature*, 2: 77–88. DOI: [10.7575/aiac.ijalel.v.2n.2p.77](#)