

# Who are ASM Journals? A Gender Analysis

Running title: A gender analysis of ASM journals

Ada K. Hagan<sup>1</sup>, Begüm D. Topçuoğlu<sup>1</sup>, Hazel Barton<sup>2</sup>, Patrick D. Schloss<sup>1†</sup>

† To whom correspondence should be addressed: [pschloss@umich.edu](mailto:pschloss@umich.edu)

1. Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI 48109

2. Department of Biology, University of Akron, Akron, OH

## Abstract

## Importance

## Introduction

Scientific societies play an integral role in the formation and maintenance of scientific communities. They host conferences that provide a forum for knowledge exchange and networking, as well as opportunities for increased visibility as a researcher. Scientific societies also frequently publish the most reputable journals in their field, facilitating the peer review process to vet new research submissions. As such, societies have great power to set both professional and scientific norms in their community by choosing what behaviors are rewarded and what types of research are accepted for publication. Authorship is a coveted measure of success in academic research as it is a key criterion for hiring and promotion processes. Accordingly, editors and reviewers of research journals have a substantial influence over the futures of hopeful authors. While the membership of scientific societies is likely to be reflective of all those who participate in the field, regardless of career track, the gatekeepers for peer review (reviewers and editors) are more reflective of the academy than the society as a whole.

Evidence has accumulated over the decades that academic research has a representation problem. While at least 50% of biology Ph.D. graduates are women, the number of women in postdoctoral positions and tenure-track positions are less than 40 and 30%, respectively (Sheltzer et al., 2014). Studies examining other metrics such as race and ethnicity find that less than 10% of all science and engineering doctorates were awarded to underrepresented minorities, while less than 25% of science and engineering doctorates in early career academia identify as non-white (NSF ADVANCE, 2014). Predictably, the disparities increase alongside academic rank, a phenomenon known as hierarchical segregation (Potvin et al., 2018). There have been many proposed reasons for these disparities (particularly against women) that include biases in training and hiring, the impact of children on career trajectories, a lack of support for primary caregivers, and a lack of recognition, which

culminate in reduced productivity as measured by research publications**Add citations.**

Recently, scientific societies and publishers have begun examining their own data to evaluate representation of, and bias against, women and minorities in their peer review processes. The American Geological Union found that while the acceptance rate of women-authored publications was greater than that for publications authored by men, women submitted fewer manuscripts than men and were used as reviewers only 20% of the time (Lerback, 2017). Fox et al., have found that for the journal *Functional Ecology*, the proportion of women invited to review depended on the gender of the editor (Fox, 2016). Despite the disproportional representation of lead women authors, several studies have concluded that there is no significant bias against papers authored by women (C&W, 2011; Fox, 2016; Handly, 2015; Edwards, 2018). Conversely, two recent studies—one of the peer review process at eLife and the other of outcomes at six ecology and evolution journals—found that women-authored papers are less likely to have positive reviews and outcomes (Murray, 2018; Fox and Paine, 2019).

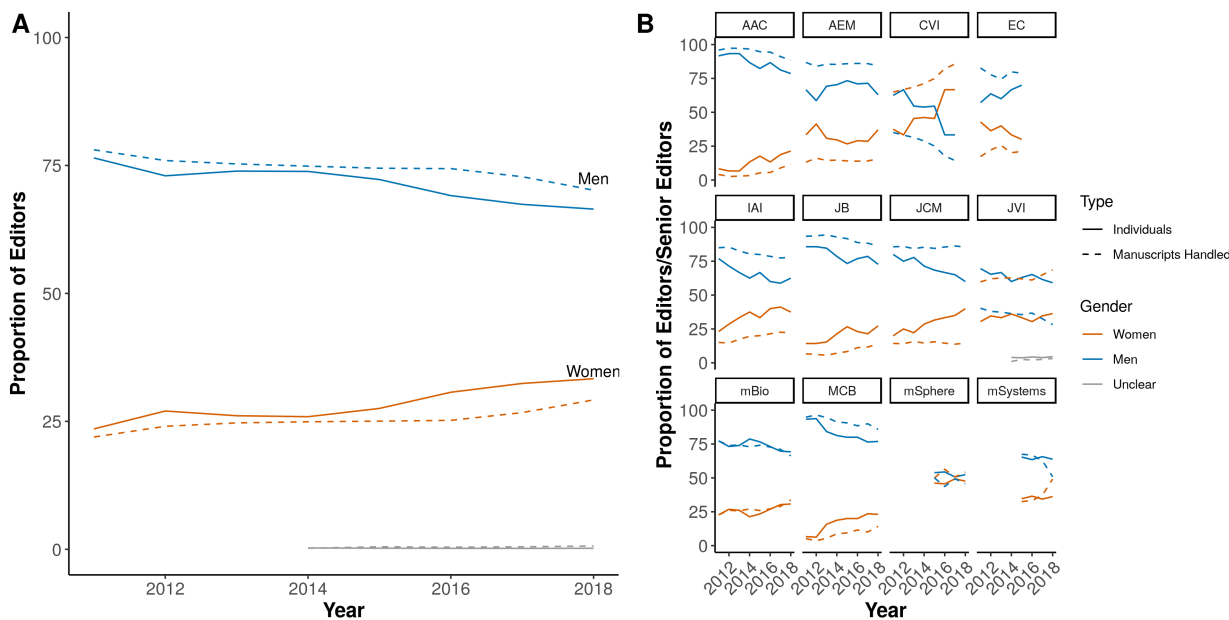
However, representation and attitudes differ by scientific field and no studies to-date seem to have investigated academic publishing in the field of microbiology. The American Society for Microbiology (ASM) is one of the largest life science societies, with an average membership of 41,000 since 1990. In its mission statement, the ASM notes that it is “an inclusive organization, engaging with and responding to the needs of its diverse constituencies” and pledges to “address all members’ needs through development and assessment of programs and services.” One of these services is the publication of microbiology research through a suite of 13 journals. Led by the ASM Journals Department, these journals boast of “quality peer review and editorial leadership.” As bastions of the microbiology field, these journals are historically responsible for the success of microbiologists. The goal of this research study is two-fold: first, to understand the representation of authors, reviewers, and editors; second, to examine the possibility of gender bias in peer review at ASM journals.

Results

This study uses the data from manuscripts submitted to ASM journals that had a final decision rendered between January 2012 and August 2018.

*Men dominate as gatekeepers and senior authors* + Editor Notes: + The proportions of men/women editors are similar for both editors and senior editors – only showed editors + There is a slow trend toward gender parity in editors + Need to add the N of each editorial group: senior editors, editors, eic – all are pooled

- + Who carries the editorial burden?
- + How many manuscripts are editors handling -- Women handle a higher median of papers than men
- + Is it proportional to the author representation? -- in some cases, but also difficult to compare
- + Does it change by EiC? -- yes, CVI/JVI have EiCs that are women



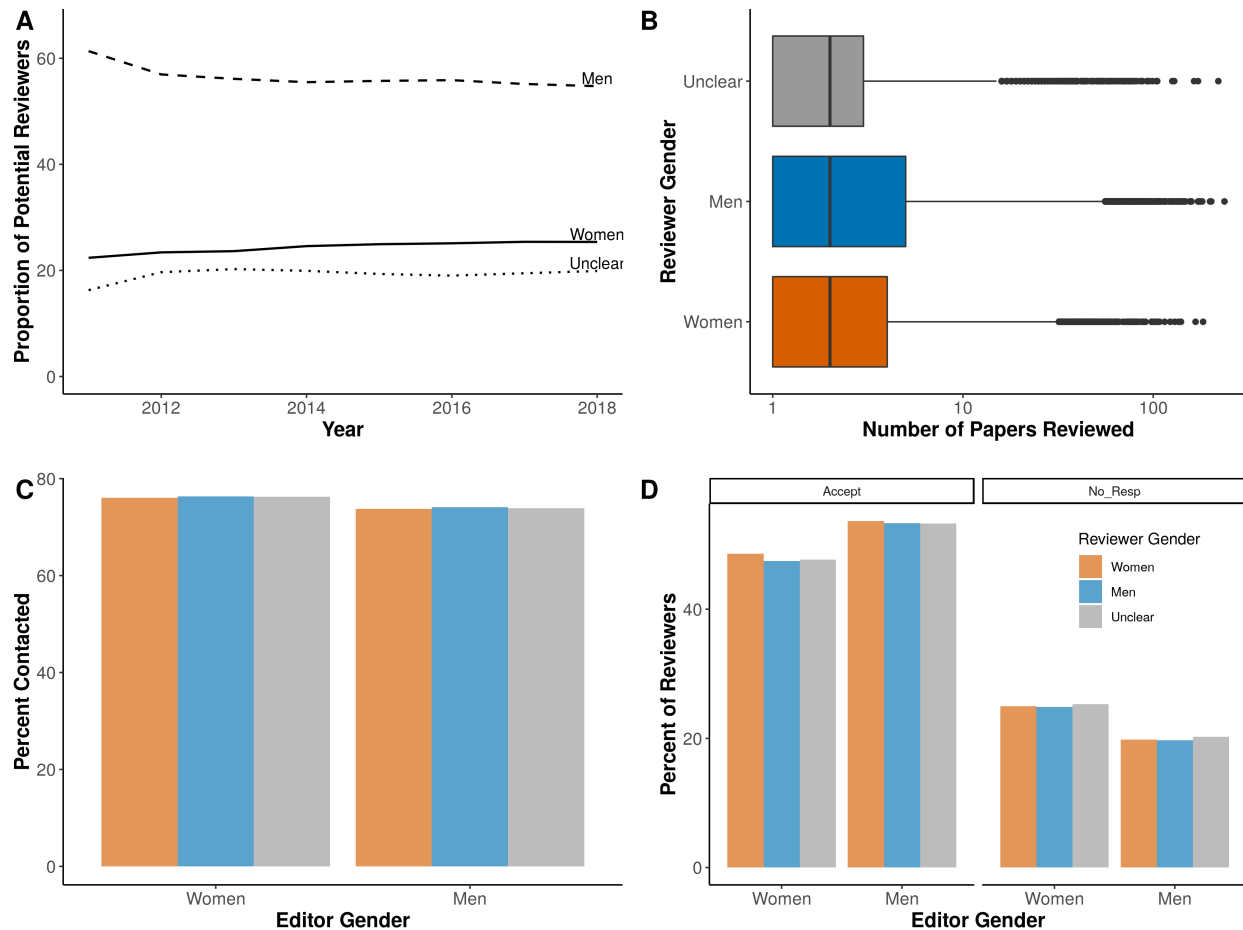
**Figure 1. Gendered representation among editors.** Proportion of editors (solid line) and their workload (dashed lines) from 2012 to 2018. Data for men are blue, and women orange. (A) All journals combined. (B) Breakdown by individual journals. Editors and senior editors are pooled together, editorial rejections are excluded. Each individual was counted once per calendar year.

68 • Reviewer Notes:

- 69 – Median number of paper reviews are equivalent across genders, though men seem  
70 slightly more likely to review more papers - trends representative of journals, except for  
71 MCB where “unclear” reviewers are a greater proportion than women
- 72 – Slight increase in the proportion of women over time (& decrease of men), more women  
73 reviewers than “unclear”
- 74 – trends and approximate proportions are equivalent for potential reviewers and actual  
75 reviewers
- 76 – greater # of women (unique individuals) have been suggested as potential reviewers  
77 than # of women senior authors – only 1/2 have actually been reviewers
- 78 – similar trend for min, except that the proportions of potential & accepted seem to be  
79 higher - calculate & add proportions relative to senior authors
- 80 – different trend for “unclear” authors where senior > potential > accepted

81 • Who carries the burden of peer review?

- 82 – How often are reviewers participating on multiple manuscripts? – calculate proportion  
83 that only reviewed one paper
- 84 – How does representation compare from potential to actual reviewer? – seems  
85 equivalent
- 86 – Is it proportional to the author representation? – yes

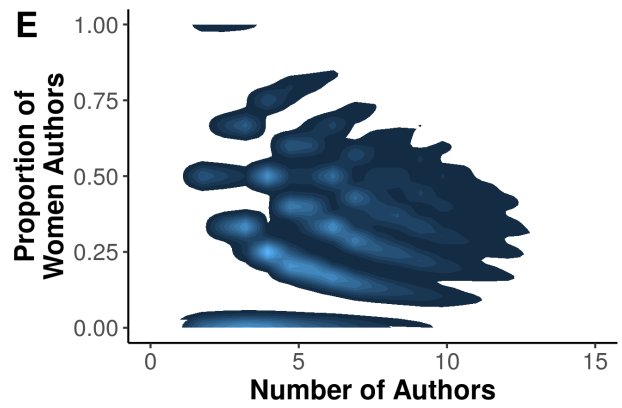
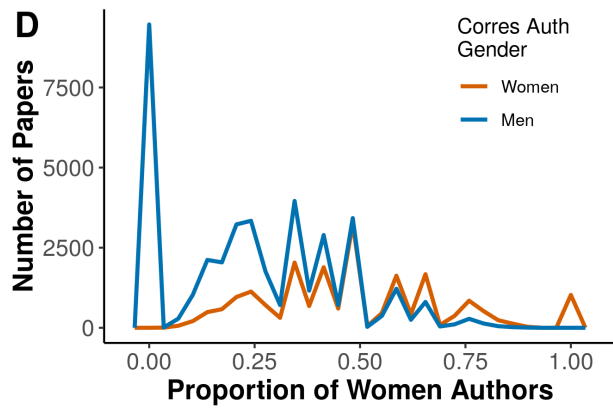
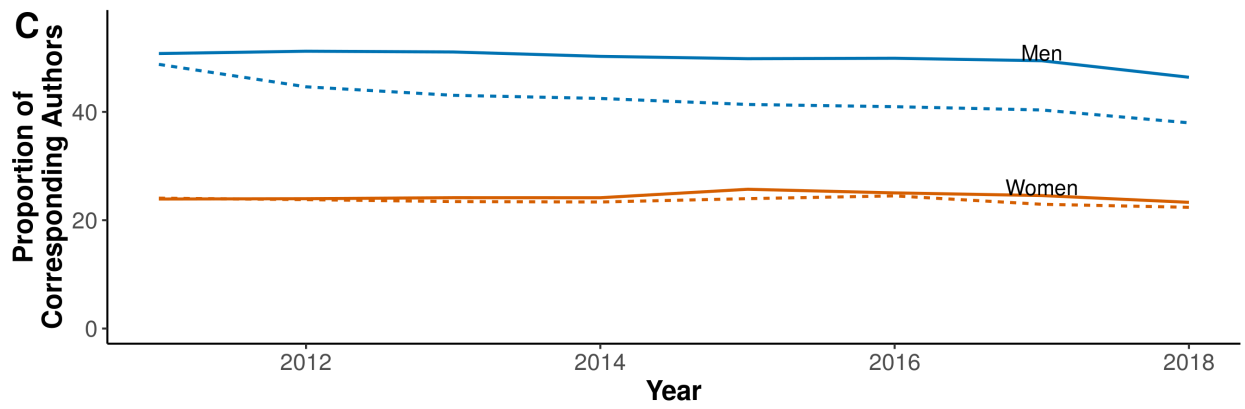
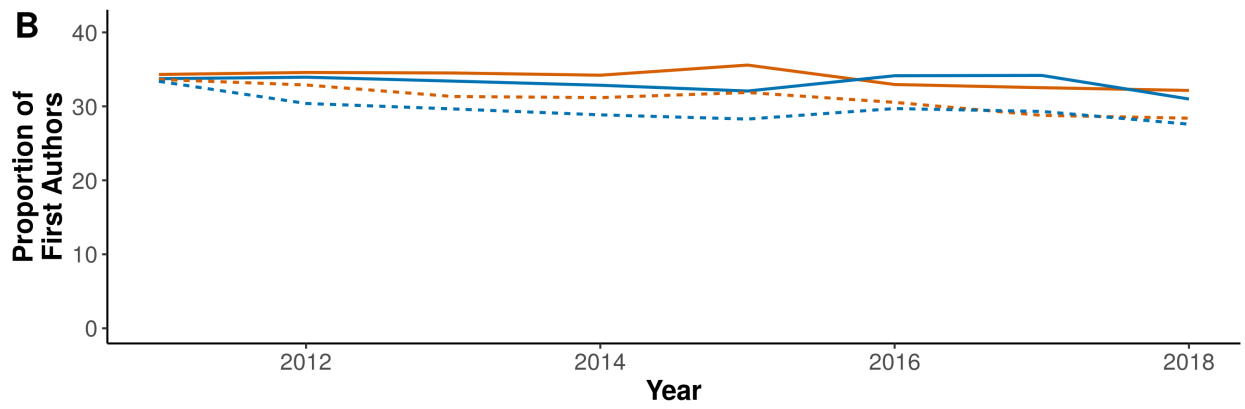
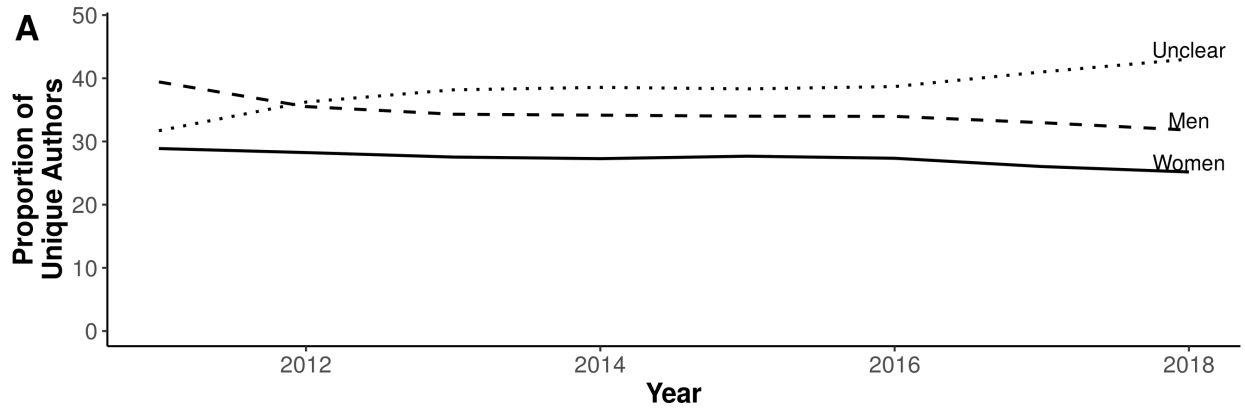


**Figure 2. Reviewer representation, workload, and response to requests to review.** (A) Proportion of each gender listed as a possible reviewer from 2012 to 2018. (B) Comparison of total papers reviewed by each individual according to gender. (C) Percent of each reviewer gender contacted to review, according to the editor's gender. (D) The percent of reviewers by gender that either accepted the opportunity to review or did not respond to a request to review, split according to the editor's gender. Reviewers were assigned one of three genders: men (blue, dashed), women (orange, solid), or unclear (gray, dotted). Each individual was counted once per calendar year.

• Notes:

- Proportions of men and women authors have decreased over time, while “unclear” has increased – gap between men & women proportions unchanged since 2012
- ratio of women:men authors is 3:4

- First & Middle author proportions are representative
- Last & corresponding author proportions are representative
- Women are first authors more often than other groups – decreasing trend of men/women first authors over time (increase of unclear)
- Trends are similar across journals & genders
- proportion of women corresponding/last authors has not increased over time
- proportion of men senior authors has decreased over time – increase of unclear senior authors
- proportion of published is higher for both men & women than submitted proportions – means that proportion of unclear published is lower than submitted
- is the increase in published proportions, proportionally equivalent for men & women?

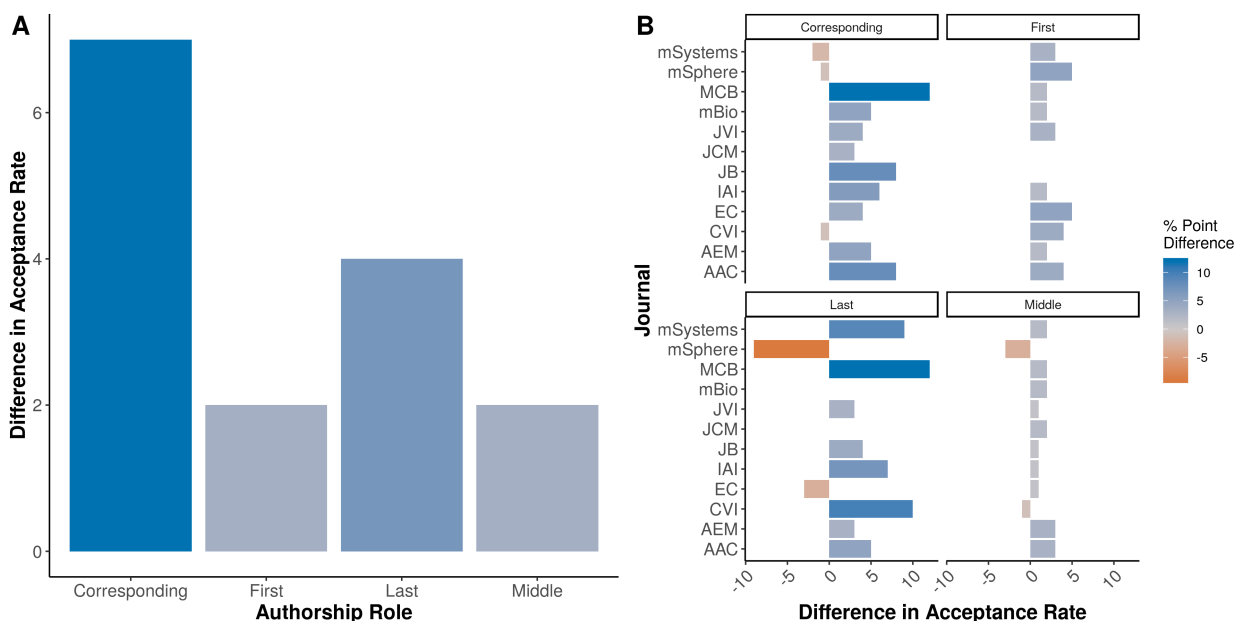




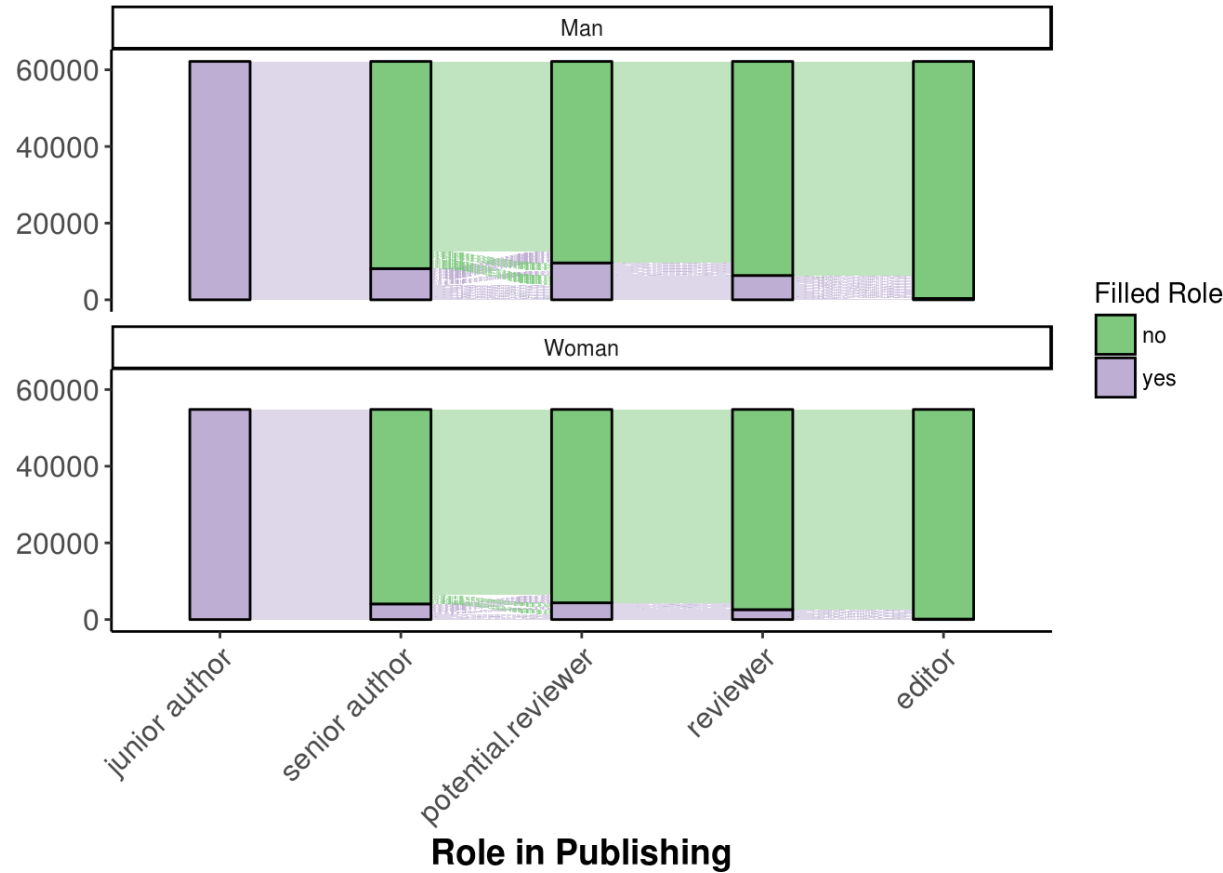
**Figure 3. Author representation by gender.** The proportion of (A) unique authors, (B) first authors, and (C) corresponding authors from 2012 - 2018. (A, B) Solid lines indicate individuals, dashed indicate proportion of manuscripts submitted. Men indicated by blue and women by orange. All individuals counted once per calendar year. The proportion of women authors on submitted papers according to (D) the gender of the corresponding author or (E) the number of authors. Unique manuscripts submitted from 2012 to 2018.

*Women are less likely to be retained through stages of the peer review system at ASM journals*

+ How does gender representation of manuscripts compare from submitted to published? + proportion of M & W authored publications seems relatively disparate from M & W authored submissions – unknown if bias + “Publication Rate Disparity” = (M publish rate) - (W publish rate) – positive == men more likely to publish (i.e., overperform); negative == women more likely to publish + Men overperform in all authorship roles across ASM journals, most clear for corresponding authors + Break it down by journals and there is clear trend to overperformance by men in both corresponding & last authorships, some exceptions + Are women being retained through peer review at the same proportion as men/unclear? + Get actual proportions at each stage + seems as if fewer women progress through each stage than men + large proportion of potential reviewers are not senior authors (both genders) + senior authors that are women are less likely to be considered as reviewers (40 vs 50%)

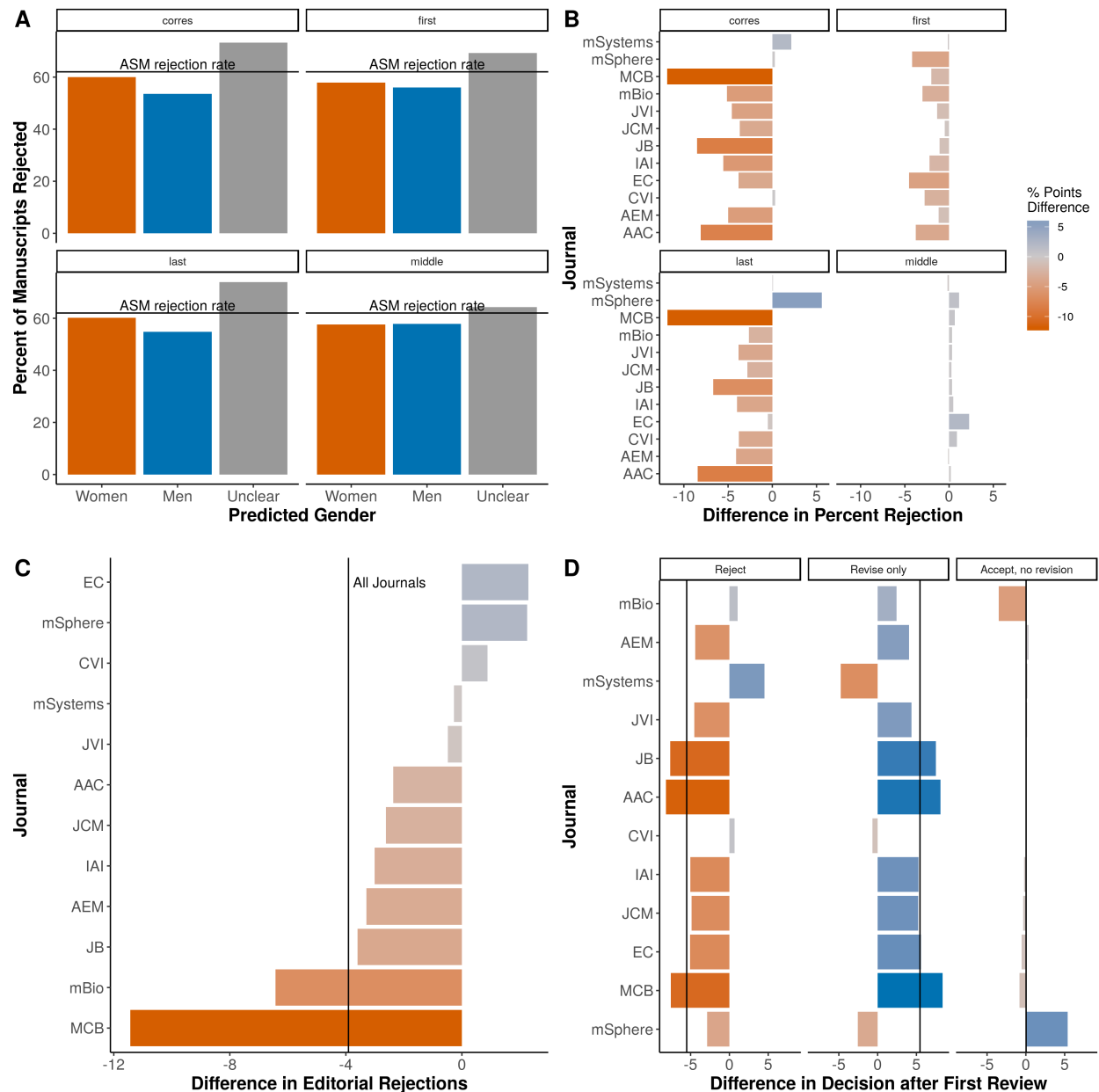


**Figure 4. The difference in percentage points of papers accepted at (A) all journals or (B) for each journal.** Unique manuscripts were split according to the gender of the corresponding, first, last, and middle author(s), and the acceptance rate for each group calculated. The difference in acceptance rate was determined by subtracting the acceptance rate of women-authored papers from men-authored papers. The shade (ranging from orange to blue) indicates the outperforming gender. No bar indicates no difference in percentage points.



**Figure 5. The retention of each gender through the publishing roles.** All junior (first or middle) authors were split by gender and tracked through their roles in academic publishing from senior author (last or corresponding), potential reviewer (considered), reviewer (accepted), or editor. Color indicates whether (purple) or not (green) the individual participated in that role at any point from 2012 to 2018.

*Papers submitted by women are rejected more frequently than those submitted by men*

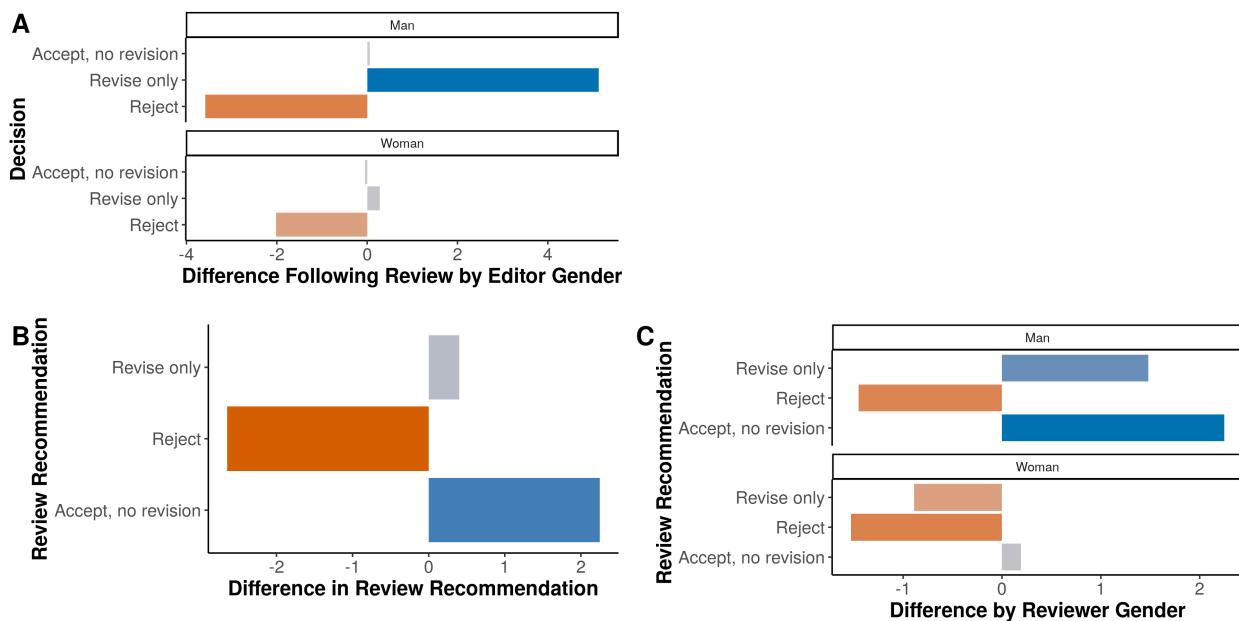


**Figure 6. Difference in rejection rates by author gender.** The percent of manuscripts rejected by author gender and type (e.g., corresponding, first, last, middle) at (A) all journals combined or at (B) each journal, which shows the difference in percent rejection rates. (C) The difference in percent editorial rejection rates at each journal, vertical line indicates the difference for all journals combined. (D) The difference in percentage points between each decision type following the first peer review, vertical lines indicate the difference value for all journals combined. The difference in rejection rates was determined by subtracting the rejection rate of women-authored papers from

men-authored papers within each category. The shade (ranging from orange to blue) indicates the outperforming gender. No bar indicates no difference in percentage points.

- try to better understand the % point difference in gender-ed performance
- NOTE: to correct for the large discrepancy in the participation of women at ASM journals during the is time period, all comparisions are made relative to the gender and population in question.
- direct comparision of the proportion of woman-authored papers that are rejected compared to men, at each author stage
  - middle/first author rejected at similar rates
  - woman-authored (corres/last) rejected more frequently than men
  - there are several journals where the overall trend is repeated and/or amplified (e.g., AAC, IAI, JB, mBio, MCB)
  - greatest effect seen for corresponding authors, use this sub-population to further examine

*Both men and women editors are more likely to reject papers authored by women than those authored by men and these preferences are partially mediated by author institution*



**Figure 7. Difference in decisions or recommendations according to the gatekeeper gender.**

(A) Effect of editor gender on the difference in percentage points for decisions following review at all journals combined. (B) Difference in percentage points for review recommendations and (C) how that is affected by reviewer gender.

**need to connect fig 7 institution info with the story – stats on gender?– SUPPLEMENTAL**

- There are two main gatekeeping roles for manuscript decisions: editors & reviewers
- This section evaluates disparities made by editors during editorial rejections, and decisions following peer review.
- Women receive more editorial rejections than men, proportionally (editorial\_rejections.R)
- may be complications due to geographic bias (eLife paper), restrict to papers submitted in US & examine institution type
- slight disparity in editorial rejection & acceptance rates according to institution type (rej\_by\_inst\_type.R)
  - greatest difference (4%) occurs for institute/medical schools
  - trend holds for most journals and is >20% diff at JCM (R2s), mBio (Federal), MCB (low research), AEM (industry research) (Supplementary\_A)
  - men from institutes/medical schools outperform women >7% for acceptance across all journals
  - 20% in favor of men at: EC (fed research), mSphere & JVI (Low research), AEM & MCB (industry research), JCM/JB/JCM (R2 institution) (Supplementary\_B)
- both men & women editors are more likely to reject women (editor\_gender\_analysis.R)
  - men editors are more likely to make revise only & accept decisions for men
  - women editors highly favor men from medical schools, slightly favor men from R2 & industry – what is the n?
  - men editors favor men from R1, R2 and medical schools, slightly favor women from low & industry research – what is the n?
- rev\_score\_analysis.R

196 – reviewers are more likely to suggest rejections for women as compared to men. No  
197 difference in revise decisions (A)

198 – men at R1, low research, medical schools & fed research are favored by reviewers over  
199 women (B)

200 • reviewer\_gender\_analysis.R

201 – both female & male reviewers are more likely to recommend rejection for women. (C)

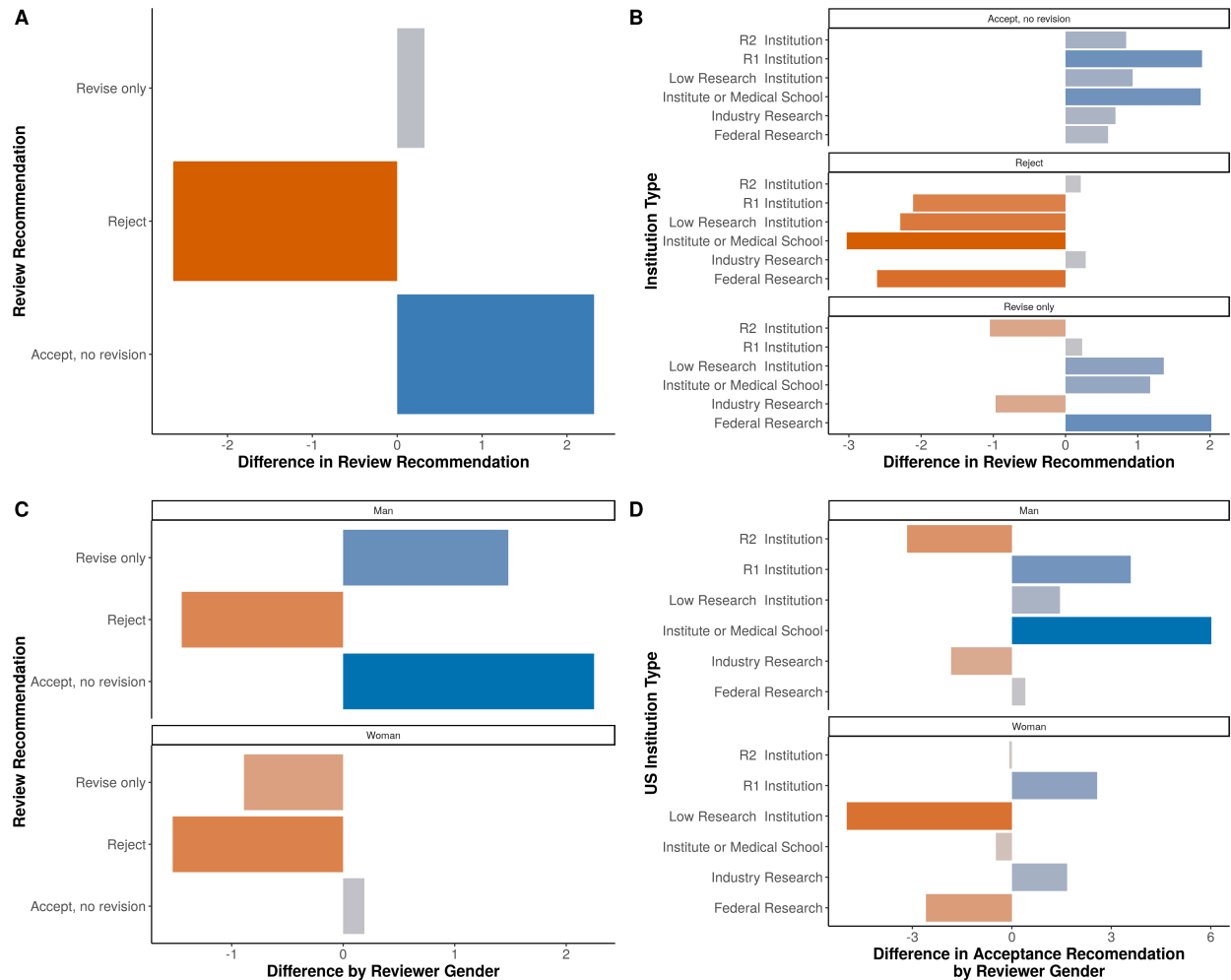
202 – Male reviewers are more likely to accept papers from men (C)

203 – women reviewers are more likely to recommend acceptance for women from low  
204 research & federal institutions (D)

205 – women reviewers are more likely to recommend acceptance for men from R1 & industry  
206 institutions (C)

207 – men are more likely to recommend acceptance for women from R2 institutions & favor  
208 men from R1 & medical institutes (D)

209 *Factors affecting competency*



**Figure 8. Comparison of time to final decision and impact by gender.** The number days (A) between when a manuscript is initially submitted then finally published and (B) that a manuscript spends in the ASM peer review system. How the impact of papers published by men (blue) versus women (orange) vary according to (C) cites and (D) total reads. Citation data includes articles published between 36 and 48 months prior to August 2018. Total reads includes both HTML and PDF online views for articles published between 12 and 24 months prior to August 2018. Impact data are divided by the number of months published.

- In addition to rejectance/acceptance rates, other disparate outcomes may occur during the peer review process
- Time in Peer Review (time\_to\_publication.R)

- Papers published by women take slightly longer (from submission to ready for publication) than men at some journals (mSphere, mBio, mSystems, CVI, JB, JCM, AEM) despite spending similar amounts of time at ASM journals (A, B)
- but do not require a greater number of revisions to be accepted than men (Supplementary\_C)
- Papers rejected following review that were submitted by women do not generally take longer (in days) to be rejected (except at mSystems) or have more revisions (Supplementary\_D & \_E)
- Decisions are returned to men & women similarly (time\_at\_journ.R)
- Peer review is not a linear process, it is cyclical and builds upon its self, may also be disparities in impact (impact\_analysis.R)
  - Women tend to receive lower cites per month published than men
  - HTML/PDF/Abstract views are equivalent
  - women are more likely to be published in lower JIF journals
  - No disparity in gendered submissions to high (mBio) vs lower (mSphere) journals (bias\_by\_jif.R)

## Discussion

- Summarize results
  - Gender
    - \* Women/men/unclear are X percent of authors, reviewers, editors
    - \* Women/men/unclear are more likely to be repeat authors/submitters
    - \* M are more likely than W to be suggested as reviewers and are/are not used as reviewers at the same proportion
    - \* M/F editors are more likely to handle multiple manuscripts – depends on journal
    - \* Women are not retained to the same extent as men/unclear
    - \* These observations do/do not correlate with gender of EiC



- \* gap between men & women performance, rejection rates
- \* women more likely to be editorially rejected or given rejection after review
- Compare to global and ASM membership stats
- Globally - microbiology researchers are 60:40, M:F - Elsevier
- ASM membership - 38.37 (sept 2018)

The under representation of women as corresponding authors in publication at ASM journals has negative consequences for their careers and microbiology. Buckley et al, suggest that being selected as a reviewer increases visibility of a researcher, which has a direct & significant impact on salary. Therefore, the underrepresentation of women as reviewers hampers their career progression and even their desire to progress since reviewing also signals adoption of the researcher into the scientific community (Buckley et al, 2014). This is supported by Lerback and Hanson who noted that “It [reviewing] provides positive feedback that a scholar is respected and participating in their field and fosters self-confidence, all of which lead to increased retention of women.” (Lerback & Hanson, 2017) Retention of women in science is important to the progress of microbiology as a field since less diversity in researchers limits the diversity of perspectives, approaches, and thus stunts the search for knowledge. In addition to boosting productivity and knowledge, more diverse and equitable organizations are more inclusive and supportive for all members (Potvin, 2018). It is thus a moral and scientific imperative for scientific societies and journals, such as ASM, to improve its own diversity, equity and inclusion efforts. The remainder of this manuscript will focus on actions that can be taken at multiple levels of the peer review system to support these efforts.

Certain attributes of biological scientific societies correlate with increased gender representation at leadership levels (Potvin, 2018). Using the scientific society “health checklist” developed from these observations, we propose the following suggestions to improve representation at society journals. First, the development of a visible mission, vision, or other commitment to equity and inclusion that includes a non-discrimination clause regarding decisions made by editors and editors-in-chief. This non-discrimination clause would be backed by a specific protocol for the reporting of, and responding to, instances of discrimination and harassment. In the long term,

society journals should begin collecting additional data about authors and gatekeepers (e.g., race, ethnicity, sexual orientation, gender identity, and disabilities). Such author data should not be readily available to journal gatekeepers, but instead kept in a disaggregated manner that allows the public presentation to track success of inclusive measures and maintain accountability. Society journals can also implement mechanisms to explicitly provide support for women and other minority groups, e.g., by providing APC waivers, reduced copyediting services, reward inclusive behavior by gatekeepers, encourage women to take up leadership positions and provide gender-neutral, non-exclusive social activities.

A common debate when filling leadership positions is whether they should be representational of the field or aspirational. For instance, since X% of corresponding authors to ASM journals are women, X% of gatekeepers of a representational leadership would be women. Conversely, 50% of gatekeepers would be women if the goal were an aspirational leadership. We argue that whether a goal should be representational or aspirational depends on the workload and visibility of the position(s). Since high visibility positions (e.g., editor, EIC) are filled by a smaller number of individuals that are responsible for recruiting more individuals into leadership, filling these positions should be done aspirationally. This allows expansion of the potential reviewer network and thus recruitment into those positions. These lower visibility positions (e.g., reviewers) require a greater number of individuals and should thus be representational of the field to avoid overburdening the minority population. Outside of leadership appointment, all parties, journals, gatekeepers, and authors, can help advance women (and other minority groups) within the peer review system. For instance, authors can suggest more women as reviewers using “Diversify” resources (e.g., DiversifyMicrobiology), while reviewers can agree to review for women editors more often. Editors can rely more on manuscript reference lists and data base searches than personal knowledge (Fox et al, 2016), and journals can improve the interactivity and functionality of the peer review selection software.

Addressing bias during peer review process is a more difficult challenge, since it is partially the result of accumulated disadvantages and microaggressions (the actions resulting from implicit biases). Implicit bias training for gatekeepers is a start, as might be double-blinded peer review, a common practice in social sciences. To support efforts of making peer review more transparent,

the review process could be unblinded following the editor's final decision on a manuscript. However, these solutions are only bandaids on a deeply infected wound since both focus on the superficial issue of individuals instead of the underlying structure of the system that has selected for the bias at hand.

Reconsidering journal scope and the overall attitude toward replicative and negative results might help address structural barriers to representation of women in peer review. Significant funds and staff are required to be competitive in highly active fields (e.g., *Clostridium difficile*, HIV), but women are often at a disadvantage for these resources. As a result, corresponding authors that are women may be spending their resources at the lesser competitive fringes of research fields. This has the disadvantage of making them seem "less competent" to those at the established center of the field. The decrease in perceived researcher competency and research validity increase the difficulty to obtain funding and publish in more traditional journals. Expanding journal scope could provide a home for these innovative research fields, bolster the field through reproducibility, and improve the competency demonstration of these researchers.

Few papers have found disparities between rejection rates of men and women and to our knowledge, this is the first paper to collectively examine this issue with either submissions data from 10+ journals or on the field of microbiology. Critics might argue that the effect size is too small to really matter or that there are too many unaccounted for factors to draw conclusions. We acknowledge that these are limitations of our study along with a limited journal dataset, an absence of reviewer comments for sentiment analysis, and that many ASM journals have a narrow focus while the broad scope journals are relatively new. All of these factors prevent us from generalizing our results across microbiology as a field. However, the consistency of the trends to benefit men corresponding authors over women, across all journals included and literature to-date confirms that this study is highly relevant for the ASM as a society and offers opportunities to address both gendered representation in microbiology and systemic barriers to peer review at our journals.

## Data and Methods

### *Data*

All manuscripts handled by ASM journals (e.g., *mBio*, *Journal of Virology*) that received an editorial decision between January 1st, 2012 and August 31st, 2018 were supplied as XML files by ASM's publishing platform, eJP. Data were extracted from the XML documents provided using R statistical software (version 3.4.4) and the XML package (R citation). Data manipulation was handled using the `tidyverse`, `lubridate`, and `xml2` packages for R. Variables of interest included: the manuscript number assigned to each submission, manuscript type (e.g., full length research, erratum, editorial), category (e.g., microbial ecology), related (previously submitted) manuscripts, versions submitted, dates (e.g., submission, decision), author data (e.g., first, last, and corresponding authorship, total number of authors), reviewer data (e.g., reviewer score, recommendation, editor decision), and person data (names, institutions, country) of the editors, authors, and reviewers.

For this analysis, only original, research-based manuscripts were included, e.g., long- and short-form research articles, New-Data Letters, Observations, Opinion/Hypothesis articles, and Fast-Track Communications.

It is common practice at ASM journals for manuscripts whose reviewers recommend extensive experimental revisions be given a decision of "reject with resubmission encouraged". If resubmitted, the authors are asked to note the previous (related) manuscript and the resubmission is assigned a new manuscript number. Multiple related manuscripts were tracked together by generating a unique grouped manuscript number based on the recorded related manuscript numbers. This grouped manuscript number served multiple purposes including: tracking a single manuscript through multiple rejections or transfers between ASM journals and to avoid duplicate counts of the same authors for the same manuscript.

Data were visualized using the `ggplot`, `scales`, `RColorBrewer`, and `ggalluvial` packages for R.

### *Bias analysis and presentation*

## *Logistic regression and clustering*

### *Gender prediction and assignment*

The gender assignment API genderize.io was used to predict an individual's gender based on their given names, and country where possible. The genderize.io platform uses data gathered from social media to predict gender based on given names with the option to include an associated language or country to enhance the odds of successful prediction. Since all manuscripts are submitted in English, precluding language association for names with special characters, names were standardized to ASCII coding (e.g., "José" to "Jose"). We next matched each individual's country against the list of X country names accepted by genderize.io. Using the GenderGuesser package for R, all unique given names associated with an accepted country were submitted to the genderize.io API and any names returned without a predictive assignment of either male or female were resubmitted without an associated country. All predictive assignments of either male or female are returned with a probability match of 0.50 or greater. The predicted genders of all given names (with and without an associated country) whose probabilities were greater or equal to our arbitrary success cut off of 0.65 were used to assign predicted gender to the individuals in our dataset. Predicted genders were assigned to individuals in the following order: first names and country, first names, middle names and country, middle names (Supplemental Figure 1). The presenting gender (man/woman) of editors and senior editors in our dataset was hand validated using Google where possible.

We recognize that biological sex (male/female) is not always equivalent to the gender that an individual presents as (man/woman), which is also distinct from the gender(s) that an individual may self-identify as. For the purposes of this manuscript, we choose to focus on the presenting gender (man/woman/unclear) based on their first names and/or appearance (for editors). In the interest of transparency, we include those individuals whose names don't allow a high degree of confidence for gender assignment in the "unclear" category of our analysis.

### *Validation of gender prediction*

We first validated the algorithm using a set of 3265 names whose gender had been hand-coded

based on appearance and were generously provided to us by \_\_\_\_ (preprint cite). The names were supplied to the genderize algorithm both with and without the accompanying country data. The data returned include the name, predicted gender (male, female, na), the probability of correct gender assignment (ranging from 0.5 to 1.0), and the number of instances the name and gender were associated together (1 or greater). The genderize algorithm returned gender predictions for 2899 when first names were given and 2167 when country data was also supplied (732 names were associated with countries unsupported by genderize).

Sensitivity and specificity, are measurements of the algorithm's tendency to return correct answers instead of false positives (e.g., a man incorrectly gendered as a woman) or false negatives (e.g., a woman incorrectly gendered as a man). The closer these values are to 1, the smaller the chance that the algorithm will return the correlating false response. Accuracy is a composite measure of the algorithm's ability to differentiate the genders correctly. These measurements were calculated from the datasets (with and without country data supplied) at three different probability threshold cutoffs: the default genderize (0.5), a probability threshold of 0.85 (0.85), and a modified probability of 0.85, which factors in the number of instances returned (pmod0.85)(citations).

At the 0.5 threshold, the dataset returned a sensitivity of 0.8943 and specificity of 0.9339 for an accuracy of 0.911, compared to a marginally higher accuracy of 0.9146 for the dataset where country data were included (Supplemental Table 1). Generally speaking, the accuracy increases as the threshold increases along with slight trade offs between sensitivity and specificity. For the purposes of our analysis, we opted to use the pmod0.85 threshold moving forward (Supplemental Table 1, in bold).

To understand the extent of geographic bias in our gender assignment against regions and languages with genderless naming conventions, or that lack social media for incorporation into the genderize algorithm, we compared the number of names predicted without associated country data to when country data was also supplied. In our test dataset, the top five countries associated with names were United States, Germany, United Kingdom, France, and China and the countries with the highest proportion of un-predicted genders when country data were supplied are Cambodia, Iceland, Indonesia, Ireland, and Mexico, where the maximum number of

names supplied ranged from 1 to 15. To determine the impact of each country towards the overall percentage of names whose genders were not predicted (27.14%), we found the difference between the percent of names unpredicted for each country and the overall percentage, multiplied by the proportion of observations from that country to the total observations and finally divided by the overall percentage of unpredicted names (Supplemental Figure 2). The top five countries with the greatest impact on unpredicted names, and thus the countries receiving the most negative bias from genderize were Canada, China, Ireland, Belgium, and Sweden (Supplemental Figure 3). These data suggest that there is likely some bias against countries with gender-neutral naming conventions (China), and indicates the stringency with which the algorithm applies gender to names that are accompanied by country data. For instance, strongly gendered names such as Peter and Pedro were not assigned gender when associated with Canada.

We next applied the genderize algorithm at the  $\text{pmod}0.85$  threshold to our journals dataset and tested its validity on a small portion. All first names collected from our dataset were submitted to genderize both with and without country data. Only those predictions whose  $\text{pmod}$  were equivalent or greater than 0.85 were carried to the next step. The predicted genders were assigned to individuals in the following order: first names and country, first names, middle names and country, middle names. Given the relatively small number of editors and senior editors in our dataset, the presenting gender (man/woman) of editors and senior editors in our dataset was hand-validated using Google where possible. Of the 1072 editor names, 938 were predicted by genderize for an accuracy of 0.9989339, thus increasing our confidence in the gender predictions where made.

In our full dataset, the five countries with the most individuals were United States, China, Japan, France, and Germany and the countries with the highest proportion of un-predicted genders were Burundi, Chad, Kingman Reef, Korea (North), Democratic People's Republic of, and Maldives, where the maximum number of names supplied ranged from 1 to 4. Proportionally, fewer names in our full dataset were assigned gender than in our validation dataset (40.01% unpredicted versus 27.14% unpredicted, respectively). Since adjusting the workflow to predict the gender of names both with and without country data, the countries receiving the most negative bias from genderize were China, Japan, Korea, Republic of, India, Taiwan, Province of China (Supplemental Figure 4). These data indicate what we previously predicted, that the genderize algorithm has bias against

440 countries with gender-neutral naming conventions.

## 441 **References**