

Gender and ethnicity trends in journal peer review: An empirical investigation using *JASIST*

Guo Zhang

School of Library and Information
Science
Indiana University Bloomington
guozhang@indiana.edu

Bradford Demarest

School of Library and Information
Science
Indiana University Bloomington
bdemares@indiana.edu

Cassidy R. Sugimoto

School of Library and Information
Science
Indiana University Bloomington
sugimoto@indiana.edu

ABSTRACT

This study analyzed rates of acceptance to the *Journal of the American Society for Information Science and Technology* between June 2009 and May 2011, in order to shed light on extra-scientific social trends (specifically, gender and ethnicity of authors and reviewers) in academic journal submission acceptance.

Keywords

Gender, ethnicity, authors, reviewers, JASIST, peer review

INTRODUCTION

Studies of extra-scientific social trends in academic journal submission acceptance have for the most part focused on bias related to social categories such as gender and ethnicity of submissions' authors or reviewers. However, it is difficult to provide an absolute "scientific" answer to questions concerning what bias exists in the peer review process. One complicating factor is whether reviewers exhibit a certain "pattern" or "emphasis" generally due to their own personal background (e.g. gender and ethnicity), exclusively in reaction to authors' backgrounds, or some combination thereof.

Such studies of extra-scientific trends in the peer review process have occurred in psychology (Lloyd, 1990) as well as medicine (Gilbert, Williams, and Lundberg, 1994) and ecology (Primack et al, 2009), but none for LIS journals. This study seeks to compare trends in author and reviewer gender and ethnicity for an LIS journal, by analyzing rates of acceptance to the *Journal of the American Society for Information Science and Technology* (hereafter *JASIST*) over a recent two-year period.

METHODS

The current study drew its sample of reviews and

recommendations from initial manuscripts submitted between June 2009 and May 2011 to *JASIST*, a single-blind peer reviewed journal (meaning reviewers know manuscript authors' names, but not vice versa). This sample contained 450 manuscripts, and 927 reviews with recommendations ("accept", "minor revision", "major revision", or "reject").

The combined list of 1393 authors and reviewers was then coded for perceived gender and perceived ethnicity, leveraging a combination of U.S. Census data and online resources ("Genealogy Data", 1990; "Baby Name Guesser", "Internet Name Database", and "NameLab"). Perceived gender and ethnicity, rather than actual gender and ethnicity based on background checks, were employed to highlight the impact of perception (specifically with regard to manuscript authors) as well as due to expediency. Perceived gender was derived from first or middle names (when first name was unavailable), and categorized as "female", "male", and "unknown". Perceived ethnicities by country were derived from last names and merged into six regional categories – Western European, East European, East Asian, South Asian, Middle East, and African, the latter five of which compose "non-Westerner".

Manuscript reviews were then divided into two sub-groups: those for single-authored papers (297 reviews) and multiple-authored papers (639 reviews) to compare proportions of recommendations across authors' and reviewers' genders and ethnicities. Due to the added complexity of multiple-authored papers, two measurements were used for this subset. The first was a measure of presence and absence of categories – for ethnicity, no Westerners vs. all Westerners, and for gender, no females vs. 1 or more females. The second measure was of categorical majorities – for ethnicity, a majority of Western authors vs. a majority of non-Westerners; for gender, a majority of female authors vs. a majority of male authors.

RESEARCH QUESTIONS AND HYPOTHESES

The following research questions were posed:

1) *Are female-authored manuscripts assigned recommendations in different proportions than male-authored manuscripts?*

2) *Do female reviewers assign different proportions of recommendations than do male reviewers?*

This is the space reserved for copyright notices.

ASIST 2012, October 28-31, 2012, Baltimore, MD, USA.
Copyright notice continues right here.

3) Are Westerner-authored manuscripts assigned recommendations in different proportions than non-Westerner authored ones?

4) Do Westerner reviewers assign different proportions of recommendations than do non-Westerners?

The expectations for the outcomes of the analysis were articulated as the following hypotheses:

H1: Female-authored manuscripts will be assigned a higher proportion of rejection and major revision, and a lower proportion of acceptance and minor revision than male-authored manuscripts (Based on Wennerås & Wold, 1997).

H2: Female reviewers will assign a higher proportion of rejection and major revision, and a lower proportion of acceptance and minor revisions than male reviewers (Based on Borsuk et al, 2009, Wing et al, 2010).

H3: Westerner-authored manuscripts will be assigned a higher proportion of acceptance and a lower proportion of rejection than non-Westerner authored ones. (Based on Bornmann, 2011; Oswald, 2008)

H4: Westerner reviewers will assign a higher proportion of acceptance and a lower proportion of rejection than non-Westerners. (Based on Marsh, Jayasinghe & Bond, 2008)

GENDER RESULTS

Tables 1 and 2 present counts and proportions by recommendation, with Table 1 reporting findings for single- and multiple-author measures and Table 2 reporting reviewer findings. Table 3 presents the statistical significance for p-values for the Chi squares test of gender results.

Rec	Single-Authored		Multiple-Authored			
	Author Gender		F (Absent/Present)		Maj	
	M	F	Absent	Present	M	F
A	8.46% (17)	12.09% (11)★	8.97% (21)★	6.92% (27)	7.85% (27)★	2.02% (2)
IR	28.86% (58)★	24.18% (22)	31.2% (73)★	28.97% (113)	32.27% (111)★	29.3% (29)
AR	33.83% (68)	36.26% (33)★	32.05% (75)	35.64% (139)★	33.72% (116)	40.4% (40)★
RE	28.86% (58)★	27.47% (25)	27.78% (65)	28.46% (111)★	26.16% (90)	28.3% (28)★
Total (N only)	201	91	234	390	344	99

Table 1. Percentages and counts by recommendation, for single- and multiple-authored paper gender measures

Rec	M	F
A	9.57% (53) ★	6.12% (20)
IR	30.69% (170) ★	25.99% (85)
MR	31.41% (174)	39.45% (129) ★
RE	28.34% (157)	28.44% (93) ★

Table 2. Percentages and counts by recommendation, by reviewer gender

Rec	Authors			Reviewers
	Single-authored	Multi-authored		
	M vs. F (N=292)	F (present/absent) (N=624)	Maj (M vs. F) (N=443)	M vs. F (N=881)
All levels (df=3)	X ² =1.51 P=0.6797	X ² =0.87 P=0.6558	X ² =5.31 P=0.1505	X ² =8.39 P=0.038675**
A vs. non-A (df=1)	X ² =0.33 P=0.3292	X ² =0.35 P=0.3519	X ² =4.27 P=0.0388**	X ² =3.22 P=0.072669*
IR vs. non-IR (df=1)	X ² =0.41 P=0.4063	X ² =0.84 P=0.5569	X ² =0.31 P=0.5748	X ² =2.20 P=0.137912
AR vs. non-AR (df=1)	X ² =0.69 P=0.6856	X ² =0.94 P=0.3604	X ² =1.51 P=0.2199	X ² =5.89 P=0.015199**
RE vs. non-RE (df=1)	X ² =0.06 P=0.6935	X ² =0.33 P=0.3313	X ² =0.18 P=0.6741	X ² =0.001 P=0.974397

Table 3. X² and p-values for the Chi squares test of gender results

NOTE: Rec-recommendations, A-acceptance, IR-minor revision, AR-major revision, RE-rejection, W-Westerner, Maj-Majority, M-male, F-female, ★ - a higher proportion in comparison to the alternate category for the same recommendation level.

* P < 0.1 ** P < 0.05

ETHNICITY RESULTS

Single-authored manuscripts. Of the 297 single-authored manuscripts, 177 (59.6%) were written by Westerners, and 114 (38.4%) by non-Westerners.

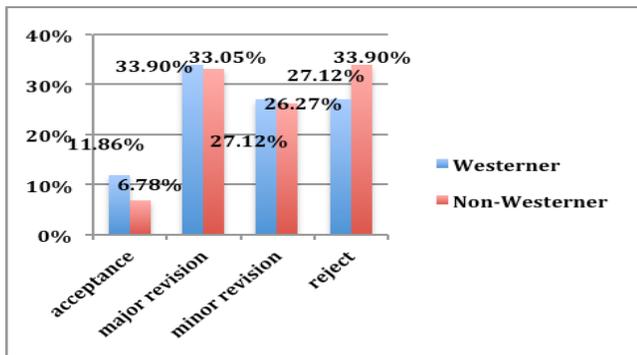


Figure 1. Proportions of different recommendation decisions across Westerner authored manuscripts and non-Westerner authored manuscripts.

Multiple-authored manuscripts. Of the 639 multiple-authored papers, 251 (or 39.28%) were written by only Western authors, while 149 (or 23.32%) were written by only non-Western authors.

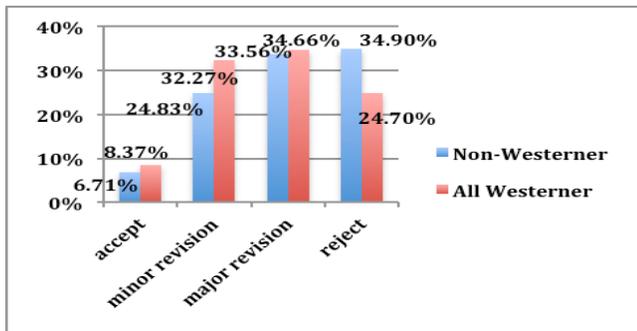


Figure 2. Proportions of different recommendation decisions across non-Westerner authored manuscripts and all Westerner authored manuscripts.

Of the 639 multiple-authored papers, 75 (11.74%) were authored by Western majorities, and 69 (10.8%) by non-Western majorities.

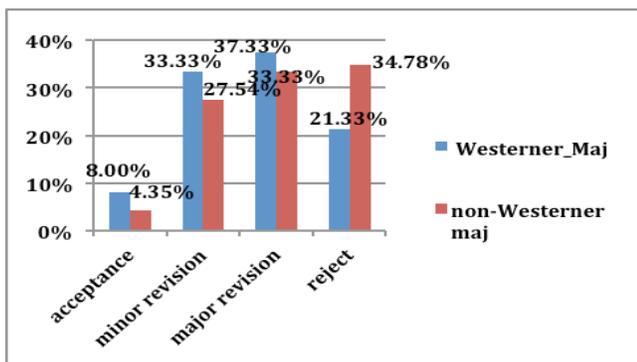


Figure 3. Proportions of recommendation decisions across Westerner-majority and non-Westerner majority authored manuscripts

Reviewers. Of the 612 reviewers, 526 (69.61%) were Westerners, while 186 (30.39%) were non-Westerners.

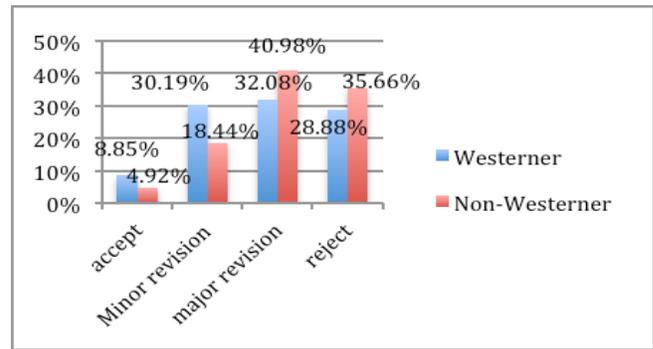


Figure 4. Proportions of recommendation levels for Westerner and non-Westerner reviewers

Table 4 presents X^2 and p-values for the Chi squares test of ethnicity results.

Rec	Authors			Reviewers
	Single-authored	Multi-authored		
	W vs. NW (N=297)	W (all/none) (N=400)	Maj (W vs. NW) (N=144)	
All levels (df=3)	$X^2=3.53$ $P=0.3168$	$X^2=5.53$ $P=0.1369$	$X^2=3.66$ $P=0.3$	$X^2=20.15$ $P=0.0002^{**}$
A vs. non-A (df=1)	$X^2=2.19$ $P=0.0602^*$	$X^2=0.36$ $P=0.5495$	$X^2=0.82$ $P=0.3657$	$X^2=3.76$ $P=0.0524^*$
IR vs. non-IR (df=1)	$X^2=0.06$ $P=0.8057$	$X^2=2.49$ $P=0.1147$	$X^2=0.57$ $P=0.4506$	$X^2=13.20$ $P=0.0003^{**}$
AR vs. non-AR (df=1)	$X^2=0.06$ $P=0.8019$	$X^2=0.05$ $P=0.822$	$X^2=0.25$ $P=0.6161$	$X^2=6.78$ $P=0.0092^{**}$
RE vs. non-RE (df=1)	$X^2=2.10$ $P=0.147$	$X^2=4.77$ $P=0.0289^*$	$X^2=3.24$ $P=0.0719^*$	$X^2=3.73$ $P=0.0533^*$

Table 4. X^2 and p-values for the Chi squares test of ethnicity results

NOTE: Rec-recommendations, A-acceptance, IR-minor revision, AR-major revision, RE-rejection, W-Westerner, NW-non-westerner, Maj-Majority

* $P < 0.1$ ** $P < 0.05$

DISCUSSION AND CONCLUSION

Of the four hypotheses, all have some support in the findings.

Author hypotheses. Mixed support occurs for H1, which hypothesizes that female authors' manuscripts are recommended for acceptance with major revision or rejection in greater proportion than are male authors' manuscripts. Single-authored papers offer no clear patterns, but multiple-authored papers with 1 or more female authors are assigned higher proportions of relatively negative recommendations (major revision, rejection) than are papers with no female authors (although Pearson's Chi Square test does not find a significant difference ($p = 0.66$). Papers with female-majority authored papers are also assigned higher proportions of major revision and rejection than are male-majority authored papers; notably, Pearson's Chi Square test does find a significant difference ($p = 0.04$) between the proportions of recommendations received by female-majority and male-majority authored papers.

H3, which hypothesizes that non-Western authors received a higher proportion of rejection and a lower proportion of acceptance than Western authors, has support as well. While minor and major revision proportions for single authors are similar between the two author groups, Western single authors have more acceptances and fewer rejections, a finding supported by Pearson's Chi Square test of recommendations for Western and non-Western single-authors ($p = 0.06$). Multiple-authored papers provide even stronger support; all-Western-authored papers having higher proportions of acceptance, minor revision, and major revision than all-non-Western-authored papers, and lower proportions of rejection. This same pattern arises between Western-majority and non-Western majority authored papers (with the latter having higher proportions of rejection, and of no other category, than the former). Particularly when comparing rejection and non-rejection decisions, all-Western-authored papers and Western majority authored papers demonstrate significant differences from their all-non-Westerner and Western-majority counterparts ($p = 0.03$ for the former, $p = 0.0719$ for the latter).

Reviewer hypotheses. H2, which asks whether female reviewers recommend higher proportions of major revision and rejection recommendations than do male reviewers, is supported by the finding that female reviewers assigned higher proportions of major revision as well as rejection than did male reviewers (although the difference in rejection proportions is a slight 0.1%), but lower proportions of acceptance and minor revision than did male reviewers. Statistical differences are demonstrated in all recommendation levels between male and female reviewers ($p = 0.04$), especially when comparing their acceptance and non-acceptance ($p = 0.072669$) and major revision and non-major-revision ($p = 0.015199$) decisions. This falls in line with studies that found female editors and female postdoctorate reviewers to be stricter than their male counterparts (e.g. Borsuk et al, 2009).

H4, which asks whether Western reviewers assign a higher proportion of acceptance and a lower proportion of

rejection than non-Western reviewers, is also supported – Western reviewers recommended higher proportions of relatively positive decisions (acceptance and minor revision) than did non-Western reviewers, and lower proportions of relatively negative decisions (major revision and rejection). Statistical differences are demonstrated in all recommendation levels between Western and non-Western reviewers ($p = 0.0002$), especially when comparing their minor revision and non-minor revision ($p = 0.0003$), major revision and non-major-revision ($p = 0.0092$) decisions. This finding is in accordance with Marsh, Jayasinghe, and Bond's (2008) report that Westerners (e.g. North American assessors) tend to give higher ratings than assessors from other countries.

Considering all four hypotheses together, and returning to one of the factors complicating studying extra-scientific social trends as generating bias in peer review, these findings suggest that a connection might exist between the experiences of female and non-Western submitting authors (vs. those of male and Western authors), and their subsequent tendencies as reviewers. As female and non-Western authors are more likely to receive major revision and rejection recommendations, they may view these recommendations as more normal for a standard level of quality in submitted work, whereas male and Western authors receiving higher frequencies of acceptance and minor revision would see these more positive recommendations as normal, and be more inclined to assign them when acting as reviewers.

LIMITATIONS AND FUTURE WORKS

While the findings may suggest this pattern, several limitations to the current study exist which curtail the certainty of such an assertion, including the size of the dataset (only 927 reviews and 1393 unique names), the number of journals analyzed (only one journal), the basis of analysis for reviewers (e.g. perceived ethnicities and gender, rather than actual collected from CVs and other background data), and ambiguity of where and how JASIST gets its pool of reviewers: JASIST reviewers, while selected by the editorial staff on the basis of subject matter expertise within the field, are not necessarily previously published authors in JASIST. Additionally, the current presentation of findings leaves aside any interactions between reviewer and submitting author categories (e.g. inter-gender vs. intra-gender effects). Further research is required to overcome these limitations, to support the theoretical foundations and to confirm empirical findings, using a larger dataset from diversified journals in LIS, as well as to interpret these differences in light of larger socio-cultural contexts. Ideally the expanded set of LIS journals would include those with double- as well as single-blind reviewing policies, thereby allowing a consideration of whether the double-blind review model might alleviate bias in peer review, or if not, whether empirical evidence might suggest that female and non-Western authors use different enough writing styles to

be recognizable and penalized even when their work is blind reviewed.

REFERENCES

- Bornmann, L. (2011). Peer Review and bibliometrics: Potentials and problems. In J. C. Shin, R. K. Toutkoushian, & U. Teichler (Hrsg.), *University rankings: theoretical basis, methodology and impacts on global higher education* (pp. 145-164). Springer: Berlin, Germany.
- Borsuk, R. M., Aarssen, L. W., Budden, A. E., Koricheva, J., Leimu, R., Tregenza, T., & Lortie, C. J. (2009). To Name or Not to Name: The Effect of Changing Author Gender on Peer Review. *BioScience*, 59(11), 985–989. doi:10.1525/bio.2009.59.11.10
- Genealogy Data: Frequently Occurring Surnames from Census 1990 – Names Files. (1990). Retrieved from http://www.census.gov/genealogy/www/data/1990surnames/names_files.html
- Gilbert, J. R., Williams, E. S., & Lundberg, G. D. (1994). Is There Gender Bias in JAMA's Peer Review Process? *JAMA: The Journal of the American Medical Association*, 272(2), 139–142. doi:10.1001/jama.1994.03520020065018
- Internet Name Database. Retrieved from <http://www.inamedb.com/name/>
- Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications: Reliability, validity, bias, and generalizability. *American Psychologist*, 63(3): 160-168
- NameLab. Retrieved from <http://genealogy.familyeducation.com/surname-origin/>
- Oswald, A. J. (2008). Can we test for bias in scientific peer-review. IZA discussion paper 3665. Bonn: Institute for the Study of Labor.
- Peters, G. "Baby Name Guesser". Retrieved from <http://www.gpeters.com/names/baby-names.php>
- Primack, R. B., Ellwood, E., Miller-Rushing, A. J., Marrs, R., & Mulligan, A. (2009). Do gender, nationality, or academic age affect review decisions? An analysis of submissions to the journal *Biological Conservation*. *Biological Conservation*, 142(11), 2415–2418. doi:10.1016/j.biocon.2009.06.021
- Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, 387 (6631): 341-343. | doi:10.1038/387341a0, 387(6631), 341–343. doi:10.1038/387341a0
- Wing, D. A., Benner, R. S., Petersen, R., Newcomb, R., & Scott, J. R. (2010). Differences in Editorial Board Reviewer Behavior Based on Gender. *Journal of Women's Health*, 19(10), 1919–1923. doi:10.1089/jwh.2009.1904