**Title: Gender Bias in Letters of Recommendation and Personal Statements for Application to Otolaryngology Residency**

**Authors: Garrett A. Berk, BS; Taylor J. Stack-Pyle, BS; Tiffany D. Ho, BS; Abdullah Zeatoun, MD; Keonho Albert Kong, MD; Mark B. Chaskes, MD, MBA; Brian D. Thorp, MD; Charles S. Ebert Jr., MD, MPH; Christine E. DeMason, MD; Brent A. Senior, MD; Adam J. Kimple, MD, PhD**

**Introduction**

Letters of Recommendation (LORs) represent an important component of medical student applications for residency. In addition to an applicant’s personal statements, they provide programs with insight into a candidate’s personality, character, work ethic, communication skills, and medical aptitude that is not otherwise apparent from the other objective components of the application (such as grades, board scores, and research experience). Results from the 2018 National Resident Matching Program (NRMP) Residency Program Director Survey highlight the importance of LORs, with only Step 1 of the United States Medical Licensing Examination (USMLE) being cited by a greater number of programs for selecting applicants to interview1. This importance is now even more significant, as beginning in 2022, USMLE Step 1 scores will be reported as Pass/Fail, making it more difficult for programs to separate applicants.

Although LORs are a crucial component of residency applications, their inherently subjective nature predispose them to potential implicit bias2. This bias, by definition, is unconscious and can skew one’s judgement in both a positive or negative direction3. Previous research has demonstrated such a bias on the basis of gender for men and women in academia, with LORs for men tending to be longer and containing more standout adjectives (such as “excellent” or “exceptional”) and descriptors pertaining to research4. Prior studies examining this phenomenon in medicine are mixed. LORs for residency applications in emergency medicine, ophthalmology, general surgery, and transplant surgery demonstrate differences in language between men and women applicants; however, the specifics regarding descriptors varied between studies5,6,7,8. An evaluation of gender-based differences in LORs to an orthopedic surgery residency program reported that language used was overall similar between men and women9. Literature analyzing linguistic differences in applicant personal statements (PSs) is markedly limited; however, gender-based differences have been noted in applications to urology, internal medicine, pediatrics, and general surgery that mirror gender stereotypes found in social psychology research10,11,12,13. To our knowledge, no publications to date have investigated gender-based differences in PSs of applicants to otolaryngology residency.

According to the *AAMC 2020 Physician Specialty Data Report*, only 18.3% of active otolaryngologists were female14. While this number has been trending upward, a lack of adequate mentors and senior role models as well reports of sexism in the surgical workplace have been reported as potential hindrances to this process15. A better understanding of potential gender bias in residency applications is therefore particularly relevant for otolaryngology, given this extreme gender imbalance in the field. In this study, we sought to analyze the linguistic differences of LORs and PSs from otolaryngology residency applications on the basis of gender.

Include other ENT studies- how are we different?

**Methods**

Study Design, Setting, and Population

This study was a retrospective analysis of all LORs and PSs for all otolaryngology residency applications submitted to an academic medical center during the 2019-20 and 2020-21 cycles (n = \*\*\*\*). The study was approved by the organization’s institutional review board. Electronic residency application system (ERAS) files were pulled from internal department archives by an administrative coordinator and applicants were assigned a unique study identification number. Demographic information (including self-reported gender, race, and ethnicity) as well as board scores and quantitative research data were abstracted. Applicant characteristics can be found in Table 1. Medical Student Performance Evaluation (MSPE) letters were not included due to their variability across medical schools.

Text Analysis

(3) Text for LORs (n = \*\*\*) and PSs (n = \*\*\*) were converted from PDF to Microsoft Word using Adobe Acrobat Pro 2020 (Adobe Systems, San Jose, California, USA). For those letters using a standardized template, the personal comments section at the end was copied for use as the narrative LOR. (2) In order to standardize word counts and linguistic analysis across LORs, only that text included between the salutation and signature was exported. (1) Quantitative analysis was performed using Linguistic Inquiry and Word Count 2015 (LIWC2015), a validated software application designed to analyze various emotional, cognitive, and structural components of written text16.

Of the 77 predefined Linguistic Inquiry and Word Count categories available in the software, 6 word categories were chosen for use in this analysis based on their relevance to applicant evaluation. Five additional word categories defined by Trix and Psenka17 (“grindstone,” “ability,” “standout,” “research,” and “teaching”) were also used. The details of the data dictionary used for this analysis are included in Supplemental Table 1. LIWC analyzes letters by counting the occurrence of words within a word category and then dividing by the number of total words in a letter to calculate the word category frequency. This frequency is then averaged over all letters to calculate the mean word category frequency. Based on the average length of the letters in our sample, a 0.\*\*\* difference in mean word frequency translates into one more word from that word category appearing when you compare three letters for one gender with three letters for the other.

Used in other studies

Specific categories

Statistical Analyses

A total of 21 variables were evaluated in this study: 10 objective and 11 subjective (see table \*\*\*). Subjective variables that the authors believe are particularly relevant to the ENT residency programs were chosen to be analyzed. [OR The subjective variables were chosen because …] Three subsets of data were analyzed: cycle 2019-20, cycle 2020-21, and two cycles combined. Within each sub-dataset, the mean differences between male and female applicants of each quality/characterization/variable were statistically assessed using t-test, when all the test’s assumptions were met, or Mann-Witney U test otherwise. Multiple testing was corrected using Benjamini-Hochberg procedure to minimize false discovery rate (FDR). [need a sentence about why Benjamini-Hochberg instead of Bonferroni?] All analyses were conducted using R 4.1.2, and p values < 0.05 were considered statistically significant.

*Table 1.*

\*\*\*

**Results**

**2020-21 F – M**

LOR:

Ability 0.79 > 0.73 p = 0.04

Clout 82.77 > 81.03 p < 0.01

Authenticity 5.18 < 6.46 p < 0.01

Achievement 3.77 > 3.61 p = 0.01

PS:

Ability 0.71 > 0.62 p = 0.01

Clout 40.43 > 38.31 p = 0.07

Tone 85.08 > 82.46 p = 0.04

Achievement 4.17 > 3.86 p < 0.01

**2019-20 F – M**

LOR:

Standout 0.66 < 0.72 p = 0.055

Grindstone 1.18 > 1.10 p = 0.04

PS: None

**Combined F – M**

LOR:

Grindstone 1.17 > 1.12 p = 0.02

Clout 82.86 > 81.94 p < 0.01

Authenticity 5.32 < 6.05 p < 0.01

Achievement 3.74 > 3.61 p < 0.01

PS:

Achievement 4.02 > 3.85 p = 0.05

**Discussion**

\*\*\*

**Conclusion**

Significant linguistic differences exist in multiple domains between LOR written for men and women applying for otolaryngology residency. The specific domains varied between application cycles, suggesting that previous studies analyzing only one year’s worth of applicants are insufficient. When combining data from two cycles, letter writers were more likely to reference achievement and use “grindstone” terminology to describe female applicants; however, letters written for male applicants were associated with a more honest, personal tone. When comparing personal statements of the applicants themselves, female applicants were more likely to reference achievement than their male counterparts.

**References**

1. Results of the 2018 NRMP Program Director Survey. <https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf>.
2. Greenwald AG, Banaji MR. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol Rev.*1995;102:4–27. doi: 10.1037/0033-295x.102.1.4.
3. Gopal DP, Chetty U, O'Donnell P, Gajria C, Blackadder-Weinstein J. Implicit bias in healthcare: clinical practice, research and decision making. *Future Healthc J*. 2021;8(1):40-48. doi:10.7861/fhj.2020-0233
4. Madera JM, Hebl MR, Martin RC. Gender and letters of recommendation for academia: agentic and communal differences. *J Appl Psychol*. 2009;94(6):1591-1599. doi:10.1037/a0016539
5. Li S, Fant AL, McCarthy DM, Miller D, Craig J, Kontrick A. Gender Differences in Language of Standardized Letter of Evaluation Narratives for Emergency Medicine Residency Applicants. *AEM Educ Train*. 2017;1(4):334-339. Published 2017 Sep 19. doi:10.1002/aet2.10057
6. Lin F, Oh SK, Gordon LK, Pineles SL, Rosenberg JB, Tsui I. Gender-based differences in letters of recommendation written for ophthalmology residency applicants. *BMC Med Educ*. 2019;19(1):476. Published 2019 Dec 30. doi:10.1186/s12909-019-1910-6
7. French JC, Zolin SJ, Lampert E, et al. Gender and Letters of Recommendation: A Linguistic Comparison of the Impact of Gender on General Surgery Residency Applicants✰. *J Surg Educ*. 2019;76(4):899-905. doi:10.1016/j.jsurg.2018.12.007
8. Hoffman A, Grant W, McCormick M, Jezewski E, Matemavi P, Langnas A. Gendered Differences in Letters of Recommendation for Transplant Surgery Fellowship Applicants. *J Surg Educ*. 2019;76(2):427-432. doi:10.1016/j.jsurg.2018.08.021
9. Kobayashi AN, Sterling RS, Tackett SA, Chee BW, Laporte DM, Humbyrd CJ. Are There Gender-based Differences in Language in Letters of Recommendation to an Orthopaedic Surgery Residency Program?. *Clin Orthop Relat Res*. 2020;478(7):1400-1408. doi:10.1097/CORR.0000000000001053
10. Demzik A, Filippou P, Chew C, et al. Gender-Based Differences in Urology Residency Applicant Personal Statements. *Urology*. 2021;150:2-8. doi:10.1016/j.urology.2020.08.066
11. Osman NY, Schonhardt-Bailey C, Walling JL, Katz JT, Alexander EK. Textual analysis of internal medicine residency personal statements: themes and gender differences. *Med Educ*. 2015;49(1):93-102. doi:10.1111/medu.12487
12. Babal JC, Gower AD, Frohna JG, Moreno MA. Linguistic analysis of pediatric residency personal statements: gender differences. *BMC Med Educ*. 2019;19(1):392. Published 2019 Oct 26. doi:10.1186/s12909-019-1838-x
13. Ostapenko L, Schonhardt-Bailey C, Sublette JW, Smink DS, Osman NY. Textual Analysis of General Surgery Residency Personal Statements: Topics and Gender Differences. *J Surg Educ*. 2018;75(3):573-581. doi:10.1016/j.jsurg.2017.09.021
14. Results of the *AAMC 2020 Physician Specialty Data Report*. <https://www.aamc.org/data-reports/workforce/interactive-data/active-physicians-largest-specialties-2019>
15. Johnson JT. Women in otolaryngology. *J Otolaryngol Head Neck Surg*. 2014;43(1):14. Published 2014 Jun 12. doi:10.1186/1916-0216-43-14
16. Tausczik YRP, Pennebaker J.W. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* *.* 2010;29:24-54.