ORIGINAL ARTICLE

Gender and Racial Bias in Radiology Residency Letters of Recommendation

Lars J. Grimm, MD, MHS[a], Rebecca A. Redmond, PhD[b], James C. Campbell, BA[c], Ashleigh S. Rosette, PhD[d]

Abstract

Objective: Perceptions of agency and communality vary by race and gender, which may be contributing to the persistent gender and racial inequality in radiology. The objective of this study was to determine if there are differences in the use of agentic and communal language in letters of recommendation for radiology residency programs based on the demographics of the applicant and letter writer.

Methods: We retrospectively reviewed letters of recommendation for 736 diagnostic radiology residency applicants to Duke University from the 2015 to 2016 interview season. We then used computerized text analysis software to calculate the frequency of agentic and communal terms and multilevel negative binominal regression to compare differences in count by applicant and letter writer demographics.

Results: We analyzed 2,624 letters of recommendation, comprising 976,489 words. The majority of applicants were male (75%, 549 of 736) and white or Asian (77%, 565 of 736). Letter writers, who were mostly male (75%, 1,979 of 2,624) and of senior rank (50%, 1,313 of 2,624), described female applicants as more agentic than men (incidence rate ratio [IRR] = 1.08, $P < .05$) and described blacks and Latinx applicants as less agentic than whites and Asians (IRR = 0.932, $P < .05$). Secondary analysis showed that female letters writers described applicants as more agentic (IRR = 1.09, $P < .05$) and more communal (IRR = 1.12, $P < .01$) than did male writers, and senior rank faculty used agentic (IRR = 0.95, $P < .05$) and communal (IRR = 0.88, $P < .01$) language less often than did junior faculty.

Conclusion: The extent to which agentic and communal language is used in letters of recommendation for diagnostic radiology residency programs differs by applicant and letter writer demographics.

Key Words: Bias, gender, letters of recommendation, race

J Am Coll Radiol 2019;∎:∎-∎. Copyright © 2019 American College of Radiology

INTRODUCTION

There is a fixed and long-standing racial and gender imbalance in diagnostic radiology, the causes of which are incompletely understood [1-3]. However, lessons may be drawn from the organizational literature that has consistently demonstrated that women and racial minorities are disadvantaged relative to men and whites, respectively [4,5]. Agency and communality are a psychological framework that describes the traditional masculine and feminine domains in Western society [6]. Communality is the female stereotype of being relationship oriented with a focus on being kind, helpful, concerned, and sympathetic, and agency refers to the male stereotype of being achievement oriented with a focus on being competent, aggressive, independent, decisive, and forceful [6-8]. A major source of disadvantage for women and racial minorities is the perceived incongruity between agentic traits and stereotypes ascribed to their subordinated groups. The communal stereotypes ascribed to women and the incompetent stereotypes (eg, inept, lazy) attributed to racial minorities are frequently perceived as incompatible with agentic characteristics that tend to define

[a]Department of Radiology, Duke University Medical Center, Durham, North Carolina.

[b]Office of Diversity and Inclusivity, Duke University, Durham, North Carolina.

[c]School of Medicine, Duke University, Durham, North Carolina.

[d]Fuqua School of Business, Duke University, Durham, North Carolina.

Corresponding author and reprints: Lars J. Grimm, MD, MHS, Duke University Medical Center, Department of Radiology, Box 3808, Durham, NC 27710; e-mail: lars.grimm@duke.edu.

The authors state that they have no conflict of interest related to the material discussed in this article.

professional roles and are deemed valuable in radiology [9-11]. Given that the field of radiology is dominated by whites and men, it is conceivable to reason that differing racial and gendered perceptions may play a role when applicants seek positions as radiologists.

Letters of recommendation are important residency selection criteria, but their subjective content lend easily to bias [12,13]. Nonradiology medical investigations have demonstrated conflicting relationships between agentic and communal language and applicant gender on residency letters of recommendation [14-17]. Ample nonmedical evidence supports the notion that women will be perceived in less agentic and more communal terms than men [11,18,19], but perceptions of women's competence, a primary dimension of agency, have increased over time [20]. Racial influences on agency and communality are even more nuanced than gender considerations. Incompetence—the antithesis of agency—is a persistent stereotype attributed to blacks and Latinx [21-24], and Asians are stereotyped as smart and hardworking, two primary dimensions of agency [25-27]. Similarly, Latinx are perceived as family-oriented (communal), yet macho (agentic); Asians are stereotyped as submissive (communal) but socially aloof (agentic); and blacks are viewed as simultaneously protective (communal) but individualistic (agentic) [23,28]. Agentic and communal perceptions are clearly influenced by gender and race, but the extent to which these perceptions vary across letters of recommendation is not fully understood. Furthermore, rater gender has been shown to influence the evaluation content [29]; hence, we also investigate its influence on the language used in recommendation letters.

We hypothesize that agentic and communal language in letters of recommendation will conform with traditional gender and racial stereotypes of applicants. Therefore, the purpose of this study is to quantify the frequency of agentic and communal language in letters of recommendation for diagnostic radiology residency applications and understand the influence of applicant and letter writer demographics.

METHODS

The Institutional Review Board from Duke University determined that this study was exempt from formal review.

Sample Population

All 736 diagnostic radiology residency applications submitted to Duke University through the Electronic Residency Application Service in the 2015 to 2016 application year were made available for analysis [30].

In line with the central purpose of the study, we first recorded the applicant's gender, applicant's self-identified race, and the letter writer gender. If the letter writer gender was not apparent from the letter of recommendation, then a clarifying Internet search was conducted. Applicant race was categorized as white or Asian or underrepresented minority (URM) or other. This was done for three reasons: (1) previous work demonstrating similarities in the evaluation of white and Asian diagnostic radiology applicants [31], (2) comparable negative perceptions of black and Latinx individuals [32], and (3) although no individual pairwise race comparisons on the use of agentic and communal terms initially emerged in our preliminary analysis, this analysis demonstrated mean similarities between white and Asian applicants and between black, Latinx, and other applicants.

Next, we included the following control variables in our analysis: (1) medical school rank categorized as top 25 by *U.S. News and World Report* 2016 rankings or not ranked in the top 25 [33-35]; (2) United States Medical Licensing Exam (USMLE) Step 1 Score; (3) an overall research productivity metric calculated by summing the number of peer-reviewed publications, poster presentations, oral presentations, and book chapters; (4) academic rank categorized as senior rank if associate or full professor and junior rank or other if medical or clinical instructor, assistant professor, or those without an academic rank, such as private practitioners; and (5) word count per letter.

Letter of Recommendation Analysis

The text of each letter of recommendation was analyzed using the Linguistic Inquiry Word Count (LIWC) software (version 1.4.0; Pennebaker Conglomerates, Inc; Austin, Texas). LIWC is a word-count-based, text analysis program that quantifies language metrics. It has been used to analyze documents in the social, physical, and medical sciences [15,29,36-38]. For this study, the entire content of the letter of recommendation was processed, excluding the greeting (eg, "Dear Committee") and closing (eg, "Sincerely, Dr Smith").

We developed a library of agentic and communal terms based on prior work establishing a library of terms [39]. In brief, the library was built based on an analysis of eight studies encompassing 3,461 data points over a 40-year period that were then included into a means-

Table 1. Description of applicants, letter writers, and letters of recommendation

Variable	All Applicants	Female Applicants	Male Applicants
Applicants, % (n)	100 (736)	25 (187)	75 (549)
Applicant race or ethnicity, % (n)			
White or Asian	77 (565)	18 (136)	58 (429)
URM or other	23 (171)	7 (51)	16 (120)
Applicant age (y), mean (SD)	27.5 (3.5)	27.6 (1.9)	27.4 (3.3)
Applicant United States Medical			
Licensing Exam Step 1 Score, mean (SD)	239.4 (15.6)	234.7 (16.9)	241.1 (14.8)
Applicant total research, mean (SD)	6.2 (9.2)	6.3 (8.8)	6.2 (9.3)
Peer-reviewed manuscripts	2.5 (5.1)	2.5 (4.6)	2.6 (5.3)
Posters	2.3 (3.3)	2.5 (3.0)	2.3 (3.4)
Oral presentations	1.2 (2.7)	1.2 (2.9)	1.2 (2.6)
Book chapters	0.1 (0.5)	0.2 (0.6)	0.1 (0.5)
Applicant medical school rank, % (n)			
Top 25 school	16 (118)	4 (30)	12 (88)
Not in top 25	84 (618)	21 (157)	63 (461)
Letter writers, % (n)	100 (2,624)	26 (674)	74 (1,950)
Letter writer gender, % (n)			
Female	25 (645)	7 (171)	18 (474)
Male	75 (1,979)	19 (503)	56 (1,476)
Letter writer academic rank, % (n)			
Junior	50 (1,311)	12 (305)	38 (1,006)
Senior	50 (1,313)	14 (369)	36 (944)
Letters of recommendation, mean (SD)	2,624	2.3 (1.1)	2.3 (1.1)
Word count per letter of recommendation, mean (SD)	364.4 (187.8)	368.7 (184.6)	362.9 (189.1)

URM = underrepresented minority.

influence-ability multifactor model followed by exploratory and then confirmatory factor analysis. From a master text file of every letter of recommendation (976,489 words), the frequency of each term was calculated. LIWC requires exact matching so modifications to the terms were required. For example, the term "ambitious" was changed to "ambitio*" to ensure both "ambition" and "ambitious" were counted. The context for each term was reviewed and removed or modified if not used in the intended manner of interest. For example, the term "dominance" was utilized in the context of "autosomal dominance," rather than describing a dominant personality, so the term was excluded from the analysis. Additional terms were added based on the context in which other agentic or communal terms were found. For example, the term "industrious" was added as an agentic descriptor because it was frequently utilized among other agentic descriptors. Edits to the library were performed in a blinded fashion by one study author (L.G.), and all additions, subtractions, or modifications were discussed by the study team before final changes were incorporated. The full library of agentic and communal terms is included in Appendix 1. Agentic

(mean: 3.50 words per letter, SD: 2.72, with average 0.97% across letters) and communal (mean: 1.42 words per letter, SD: 1.56, with average 0.39% across letters) terms that were utilized at least 200 times were included in our analysis.

Statistical Analysis

Agentic and communal letter content was compared across applicant and writer characteristics using Kruskal-Wallis tests and Dunn's post hoc pairwise comparisons with a Bonferroni correction for multiple groups. Next, multilevel negative binomial regression (NBR) (Stata command menbreg) was used for modeling the counts of agentic and communal terms. Multilevel NBR was used because there were multiple observations per applicant, the outcomes were counts, and overdispersion (variance > mean) characterizes each outcome's distribution. NBR includes a parameter (alpha) along with an error term to account for extra dispersion. Random effects were modeled at the subject level, with fixed effects estimated for the independent variables and controls. Two models are reported for each outcome. First, we model the outcomes conditioned on control

Table 2. Agentic and communal terms by applicant and letter writer characteristics, unadjusted for covariates

Characteristic	No. of Letters	Agency		Communality	
		Mean Count (SD)	Mean % of Letter (SD)	Mean Count (SD)	Mean % of Letter (SD)
Overall	2,624	3.50 (2.72)	0.97 (067)	1.42 (1.56)	0.39 (0.42)
Applicant gender					
Female	674	3.70 (2.87)	0.99 (0.66)	1.52 (1.72)	0.40 (0.45)
Male	1,950	3.44 (2.66)	0.96 (0.68)	1.38 (1.50)	0.38 (0.40)
Applicant race or ethnicity					
White or Asian	2,011	3.59 (2.74)	0.98 (0.67)	1.42 (1.52)	0.38 (0.40)
URM or other	613	3.24 (2.62)	0.93 (0.67)	1.42 (1.69)	0.41 (0.46)
Letter writer gender					
Female	645	3.65 (2.74)	1.00 (0.71)	1.59 (1.71)	0.42 (0.44)
Male	1,979	3.46 (2.71)	0.95 (0.66)	1.36 (1.50)	0.37 (0.41)
Letter writer rank					
Junior	1,311	3.37 (2.55)	1.01 (0.70)	1.42 (1.53)	0.41 (0.43)
Senior	1,313	3.64 (2.86)	0.93 (0.63)	1.41 (1.59)	0.34 (0.40)
Female writer × female applicant	171	3.54 (2.49)	0.99 (0.66)	1.74 (1.89)	0.47 (0.51)
Female writer × male applicant	474	3.69 (2.83)	1.00 (0.72)	1.54 (1.64)	0.41 (0.41)
Male writer × female applicant	503	3.76 (2.99)	1.00 (0.65)	1.45 (1.65)	0.38 (0.43)
Male writer × male applicant	1,476	3.36 (2.60)	.94 (0.66)	1.33 (1.45)	0.37 (0.40)

Mean count (standard deviation) and mean percent of agentic and communal terms per letter. Unadjusted for covariates and subject-level clustering. URM = underrepresented minority.

variables. Second, we model the main effects of applicant gender, race, and writer gender with controls, along with an interaction term between applicant and writer gender.

Tests for outliers involved inspecting cases with extreme values on the outcomes (eg, counts greater than 12 for agency and 9 for communality) and conducting iterative regression models removing these values to test sensitivity of results. Because the results were unchanged by excluding these values, we used the full data for our analysis. Incidence rate ratios (IRRs), which provide the effect size, were reported. All statistical analyses were conducted using Stata/SE by RR (version 15; StataCorp LLC; College Station, Texas).

RESULTS

We analyzed 2,624 letters of recommendation for 736 applicants, comprising 976,489 words. Table 1 describes the applicants, letter writers, and letters of recommendation. In brief, applicants were mostly male (75%, 549 of 736), white or Asian (77%, 565 of 736), and not graduating from a top 25 medical school (84%, 618 of 736). The applicant mean age was 27.5 ± 3.5 years, USMLE Step 1 score was 239.4, and research metric was 6.2 (sum of peer-reviewed publications, posters, oral presentations, and book chapters).

The only demographic difference between male and female applicants was a higher USMLE Step 1 score for male applicants (241.1 versus 234.5, $P < .001$). Letter writers were most commonly male (75%, 1,979 of 2,624) and senior rank (50%, 1,313 of 2,624). As a result, letters of recommendation were most commonly written by male letter writers for male applicants (56%, 1,476 of 2,624).

The mean counts and mean percent of agentic and communal terms by applicant and letter writer demographics, as well as for applicant and letter writer gender pairings, are shown in Table 2. The results from the multilevel NBR model for agentic and communal language with main effects and the applicant × writer gender interaction are shown in Table 3 and Figure 1. Letter writers used more agentic terms when describing female applicants as opposed to male applicants (IRR = 1.08, Model 2 in Table 3, $P < .05$). Exploratory analysis demonstrates that this finding was primarily driven by differences in the use of three specific agentic terms (mean count differences ≥ 0.03 per letter for skill, lead, and dedication), indicating the influence of specific types of agentic perceptions (Table 4). Additionally, letter writers used fewer agentic terms when describing URMs or other applicants as compared with white or Asian applicants (IRR = 0.93, $P < .05$). Exploratory analysis showed that six agentic terms

Table 3. Regression results: multilevel negative binomial regression of agentic and communal terms in letters of recommendation by applicant and letter writer characteristics

Variable	Outcome: Agency Terms		Outcome: Communality Terms	
	Model 1	Model 2	Model 3	Model 4
Applicant female (Ref: applicant male)		1.08* (0.04)		1.04 (0.06)
Applicant URM or other (Ref: applicant white or Asian)		0.93* (0.03)		1.03 (0.05)
Writer female (Ref: writer male)		1.09* (.04)		1.12* (0.06)
Applicant gender × writer gender		0.90 (0.06)		1.07 (0.11)
Writer senior rank (Ref: writer junior or other rank)	0.95* (0.03)	0.95* (0.03)	0.87† (0.04)	0.88‡ (0.04)
Top 25 medical school (Ref: below top 25)	0.94 (0.04)	0.94 (0.04)	1.06 (0.06)	1.04 (0.06)
Step 1 score	0.98 (0.01)	0.98 (0.01)	0.96* (0.02)	0.97 (0.02)
Total research	1.00 (0.00)	1.00 (0.00)	0.99† (0.00)	0.99† (0.00)
Word count	1.00† (0.00)	1.00† (0.00)	1.00† (0.00)	1.00† (0.00)
Intercept	1.76† (0.06)	1.74† (0.06)	0.72† (0.03)	0.69† (0.04)
Random effects: applicant level	1.02‡ (0.01)	1.02‡ (0.01)	1.02 (0.02)	1.02 (0.02)
Likelihood ratio test	9.54‡	8.62‡	1.31§	1.06
AIC	11,165	11,161	8,137	8,135

In all, there were 2,624 letters for 736 applicants. Multilevel negative binomial regression with subject level clustering. Incidence rate ratio (IRR) and standard error (SE) reported. IRR > 1.00 indicates increasing rate of terms used; IRR < 1.00 indicates decreasing rate of terms used. AIC = Akaike information criterion, URM = under-represented minority.

*$P < .05$.
†$P < .001$.
‡$P < .01$.
§$P < .10$.

(mean count differences ≥ 0.03 per letter for skill, strong, lead, confidence, active, and work ethic) mostly influenced this difference, indicating a more generalized difference in perception when compared with the noted gender differences (Table 4). For communality, no differences were noted based on the gender and race of the applicant (Model 4 in Table 3).

Our analysis of the influence of the demographics of the letter writer showed that women used agentic terms more frequently (IRR = 1.09, $P < .05$) and communal terms more frequently (IRR = 1.12, $P < .05$) than did their male counterparts when describing applicants. Furthermore, senior-rank letter writers used agentic terms less frequently (IRR = 0.95, $P < .05$) and communal terms less frequently (IRR = 0.88, $P < .01$) than did their junior ranked counterparts. Of note, the mean letter length was 407 words for senior rank letter writers and 344 words for junior rank letter writers, which compensates for the slightly higher mean counts for senior rank letter writers as shown in Table 2. Neither letter writer gender nor letter writer rank qualified the differences in the use of agentic content noted between female and male applicants and between whites or Asians and URMs or other applicants. Although there were trends in the use of agentic and communal language based on the applicant gender × writer gender, they were not statistically significant (Table 3).

DISCUSSION

Our results show that letter writers used agentic language more frequently to describe female applicants than male applicants. At first glance, this result may seem surprising because perceptions of agency seem in conflict with traditional feminine stereotypes [11,40]. However, these findings are consistent with recent research demonstrating that women are sometimes perceived as possessing more rather than less agency when specific types of agency (ie, competence) are under consideration [11,41]. For example, top women leaders have been shown to be evaluated as more agentic than male leaders, and these findings were explained by differing perceptions in double standards of competence [6]. Also, women leaders were evaluated more favorably when enacting self-reliance rather than dominance [41]. Our results are consistent with these findings because our exploratory post hoc analysis showed that differences in terms like "skill" and "dedication," rather than terms like "dominance" and "tough," accounted for higher descriptors of agency for women when compared with men. Hence, the extent to which women were described more favorably than men is contingent on the type of agency under consideration. Although beyond the scope of this project, the
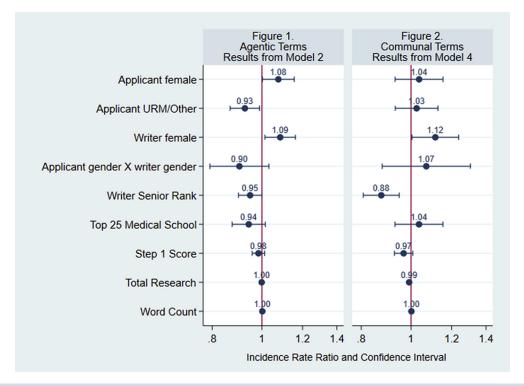
Fig 1. Forest plots of predictor variables for agentic (Figure 1) and communal (Figure 2) terms based on Models 2 and 4 as shown in Table 3. Incidence rate ratios and 95% confidence intervals are shown.

utilization of agentic and communal language for men and women may be influenced by the specialty of medicine, with certain personality characteristics deemed more or less favorable in different fields.

When considering race, letter writers were more inclined to use agency when describing white and Asian applicants rather than black and Latinx applicants. Furthermore, when compared with the gender differences, letter writers used a more inclusive set of agentic descriptors, like work ethic, confidence, and leader potential, to distinguish whites and Asians from URMs. This suggests URMs may be perceived to differ from whites and Asians on more types of agency than women are perceived to differ from men. Although racial differences in the field of radiology have not been well studied, these findings are in line with generally favorable biases toward whites and negative perceptions that have been shown to persist toward racial minorities in social science research [21-24] and in the limited research that has considered race in medical school admissions [42,43]. For example, when utilizing the implicit association test, medical doctors—including medical school admissions committee members—have been shown to have a pro-white bias [42,43]. However, it is important to note that racial minorities are less inclined to exhibit pro-white biases

on the Implicit Association Test [44], and we were not able to attain the race of the letter writers for this study.

Although not the primary focus of our analysis, we also found that the frequency of agentic and communal language differed by two demographic characteristics of the letter writer: gender and rank. Specifically, women writers used agentic and communal language more frequently than men writers. This is slightly at odds with an analysis of standardized dean's letters by Isaac et al, which did not find differences in language based on letter writer gender but rather based on the letter writer–applicant gender combination [29]. Future research should attempt to explore and reconcile the distinctions between these studies. When considering faculty rank, junior faculty used communal and agentic descriptors at a greater rate than did senior faculty (ie, slightly lower mean counts but shorter letters for junior faculty). This finding may reflect differences in experience because senior faculty who have a larger pool of applicants for comparison may be less effusive in their descriptions or rely on more templated language copied from prior letters. Regardless, these results indicate that the demographics of the letter writer do have an influence on the language used and programs may need to initiate training programs for letter writers.

Table 4. Agentic and communal terms used at least 200 times by applicant characteristic

Variable	Overall	Male Applicant (n = 549)	Female Applicant (n = 187)	White or Asian Applicant (n = 565)	URM or Other applicant (n = 171)
No. of letters	2,624	1,950	674	2,011	613
Agency (overall)	3.50 (2.72)	3.44 (2.66)	3.7 (2.87)	3.59 (2.74)	3.24 (2.62)
Skill	0.83 (1.16)	0.82 (1.15)	.88 (1.17)	0.86 (1.18)	0.74 (1.08)
Strong	0.50 (0.81)	0.50 (0.80)	0.51 (0.85)	0.51 (0.81)	0.47 (0.83)
Lead	0.36 (0.86)	0.33 (0.82)	0.43 (0.98)	0.38 (0.88)	0.30 (0.80)
Dedication	0.31 (0.64)	0.30 (0.62)	0.34 (0.71)	0.31 (0.63)	0.32 (0.66)
Confidence	0.31 (0.59)	0.31 (0.57)	0.33 (0.62)	0.32 (0.59)	0.28 (0.58)
Active	0.26 (0.57)	0.26 (0.58)	0.26 (0.55)	0.27 (0.58)	0.23 (0.51)
Intelligent	0.16 (0.39)	0.16 (0.39)	0.18 (0.40)	0.16 (0.39)	0.17 (0.40)
Work ethic	0.25 (0.51)	0.25 (0.51)	0.24 (0.49)	0.26 (0.51)	0.23 (0.49)
Hard working	0.16 (0.41)	0.15 (0.40)	0.16 (0.41)	0.16 (0.41)	0.14 (0.38)
Desire	0.11 (0.38)	0.11 (0.38)	0.11 (0.38)	0.11 (0.38)	0.11 (0.37)
Independent	0.12 (0.39)	0.12 (0.39)	0.13 (0.39)	0.13 (0.40)	0.12 (0.37)
Achieve	0.12 (0.39)	0.12 (0.39)	0.12 (0.38)	0.12 (0.38)	0.14 (0.39)
Communality (overall)	1.42 (1.56)	1.38 (1.50)	1.52 (1.72)	1.42 (1.52)	1.42 (1.69)
Help	0.47 (0.84)	0.46 (0.81)	0.52 (0.91)	0.47 (0.82)	0.48 (0.91)
Eager	0.21 (0.49)	0.21 (0.50)	0.21 (0.48)	0.21 (0.49)	0.22 (0.50)
Compassion	0.20 (0.49)	0.20 (0.49)	0.21 (0.48)	0.20 (0.50)	0.20 (0.44)
Happy	0.13 (0.39)	0.13 (0.38)	0.14 (0.39)	0.14 (0.39)	0.12 (0.37)
Kind	0.13 (0.38)	0.12 (0.38)	0.14 (0.39)	0.13 (0.39)	0.11 (0.36)
Caring	0.10 (0.33)	0.09 (0.32)	0.12 (0.38)	0.10 (0.32)	0.11 (0.37)
Friend	0.09 (0.31)	0.08 (0.30)	0.09 (0.34)	0.08 (0.31)	0.09 (0.34)

Results are mean count (SD) of agentic and communal terms per letter.

There are some limitations to this study. First, we included all applicants who applied to Duke University, but this represents only a portion (54%, 736 of 1,360) of all diagnostic radiology applicants [45]. Second, we only had sparse information on the letter writer (gender, academic rank). Writers likely vary a great deal based on skill, experience, and training, for which we could not control. Third, other factors that influence application success were not included on the application, such as physical appearance, for which we could not control [31]). Future research should address these limitations in follow-up studies.

In conclusion, we demonstrated that agentic language was more frequently used to describe female applicants and less frequently used to describe URM or other applicants. Secondary analysis demonstrated that female and junior rank letter writers use both agentic and communal language at a greater rate compared with their male and senior rank counterparts. Further work is justified to understand the implications of agentic and communal language on applicant match success and to develop better practice guidelines or educational interventions for letter writers.

TAKE-HOME POINTS

- In diagnostic radiology residency letters of recommendation, agentic language was used more frequently to describe female applicants (IRR = 1.08, $P < .05$) and less frequently to describe black and Latinx applicants (IRR = 0.93, $P < .05$).
- Female letter writers use agentic (IRR = 1.09, $P < .05$) and communal (IRR = 1.12, $P < .01$) language at a greater rate than male writers.
- Senior rank faculty letter writers use agentic (IRR = 0.95, $P < .05$) and communal (IRR = 0.88, $P < .01$) language at a lower rate than junior rank faculty.
- Medical schools may wish to implement training programs for faculty who serve as medical student mentors who will be asked to write letters of recommendation.
- Residency selection committees should be mindful of the differences in language on letters of recommendation when making decisions regarding interview invitations and residency rank list positioning.

ADDITIONAL RESOURCES

Additional resources can be found online at: https://doi.org/10.1016/j.jacr.2019.08.008.

REFERENCES

1. Cater SW, Yoon SC, Lowell DA, et al. Bridging the gap: identifying global trends in gender disparity among the radiology physician workforce. Acad Radiol 2018;25:1052-61.

2. Campbell JC, Yoon SC, Cater SW, Grimm LJ. Factors influencing the gender breakdown of academic radiology residency programs. J Am Coll Radiol 2017;14:958-62.

3. Grimm LJ, Ngo J, Pisano ED, Yoon S. Men (and women) in academic radiology: how can we reduce the gender discrepancy? AJR Am J Roentgenol 2016;206:678-80.

4. Eagly AH. Female leadership advantage and disadvantage: resolving the contradictions. Psychol Women Q 2007;31:1-12.

5. Fiske ST, Lee TL. *Stereotypes and prejudice create workplace discrimination. Diversity at work.* United Kingdom; Cambridge University Press: 2008:13-52.

6. Rosette AS, Tost LP. Agentic women and communal leadership: how role prescriptions confer advantage to top women leaders. J Appl Psychol 2010;95:221-35.

7. Abele AE. The dynamics of masculine-agentic and feminine-communal traits: findings from a prospective study. J Pers Soc Psychol 2003;85:768-76.

8. Bakan D. The duality of human existence: an essay on psychology and religion. Chicago, IL: Rand McNally; 1966.

9. National Residency Match Program. Results of the 2018 NRMP Program Director Survey. Available at: https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf. Accessed January 9, 2019.

10. Rosette AS, Leonardelli GJ, Phillips KW. The white standard: racial bias in leader categorization. J Appl Psychol 2008;93:758-77.

11. Eagly AH, Karau SJ. Role congruity theory of prejudice toward female leaders. Psychol Rev 2002;109:573-98.

12. Yao Y, Okada Y, Yamato M, Ohtomo K. Communicating vein between the left renal vein and left ascending lumber vein: incidence and significance on abdominal CT. Radiat Med 2003;21:252-7.

13. National Academy of Sciences (US), National Academy of Engineering (US), and Institute of Medicine (US) Committee on Maximizing the Potential of Women in Academic Science and Engineering. Beyond bias and barriers: fulfilling the potential of women in academic science and engineering. Washington DC 2007.

14. 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 2018;76:899-905.

15. 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;1:334-9.

16. 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 2018;76:427-32.

17. Messner AH, Shimahara E. Letters of recommendation to an otolaryngology/head and neck surgery residency program: their function and the role of gender. Laryngoscope 2008;118:1335-44.

18. Schein VE. The relationship between sex role stereotypes and requisite management characteristics. J Appl Psychol 1973;57:95-100.

19. Fiske ST, Cuddy AJ, Glick P, Xu J. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. J Pers Soc Psychol 2002;82:878-902.

20. Eagly AH, Nater C, Miller D, Kaumann M, Sczesny S. Gender stereotypes have changed: a cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. Am Psychol 2019 Jul 18. https://doi.org/10.1037/amp0000494. [Epub ahead of print].

21. Katz D, Braly K. Racial stereotypes of one hundred college students. J Abnorm Soc Psychol 1933;28:280-90.

22. Niemann YF, Jennings L, Rozelle RM, Baxter JC, Sullivan E. Use of free responses and cluster analysis to determine stereotypes of eight groups. Pers Soc Psychol Bull 1994;20:379-90.

23. Ghavami N, Peplau LA. An intersectional analysis of gender and ethnic stereotypes: testing three hypotheses. Psychol Women Q 2013;37:113-27.

24. Devine PG, Elliot AJ. Are racial stereotypes really fading? The Princeton Trilogy revisited. Pers Soc Psychol Bull 1995;21:1139-50.

25. Bell MP, Harrison DA, McLaughlin ME. Asian American attitudes toward affirmative action in employment: implications for the model minority myth. J Appl Behav Sci 1997;33:356-77.

26. Kao G. Asian Americans as model minorities? A look at their academic performance. Am J Educ 1995;103:121-59.

27. Wood PS, Altmaier EM, Franken EA Jr, Evens RG, Schlechte JA. Factors influencing choice of academic or practice careers in radiology. Invest Radiol 1990;25:675-7.

28. Rosette AS, Koval CZ, Ma A, Livingston R. Race matters for women leaders: intersectional effects on agentic deficiencies and penalties. Leadersh Q 2016;27:429-45.

29. Isaac C, Chertoff J, Lee B, Carnes M. Do students' and authors' genders affect evaluations? A linguistic analysis of medical student performance evaluations. Acad Med 2011;86:59-66.

30. Kurmis AP, Kurmis TP. Exploring the relationship between impact factor and manuscript rejection rates in radiologic journals. Acad Radiol 2006;13:77-83.

31. Maxfield CM, Thorpe MP, Desser TS, et al. Appearance-based discrimination in radiology resident selection. Paper presented at: Radiological Society of North America Annual Meeting; Nov 25-30 2018; Chicago, IL.

32. Priest N, Slopen N, Woolford S, et al. Correction: stereotyping across intersections of race and age: racial stereotyping among white adults working with children. PLoS One 2018;13:e0205614.

33. Wayback Machine. Internet Archive. Available at: https://archive.org/web/. Accessed January 22, 2019.

34. U.S. News & World Report. Best medical schools: research. Available at: https://www.usnews.com/best-graduate-schools/top-medical-schools. Accessed January 22, 2019.

35. Grimm LJ, Shapiro LM, Singhapricha T, Mazurowski MA, Desser TS, Maxfield CM. Predictors of an academic career on radiology residency applications. Acad Radiol 2014;21:685-90.

36. Kahn JH, Tobin RM, Massey AE, Anderson JA. Measuring emotional expression with the linguistic inquiry and word count. Am J Psychol 2007;120:263-86.

37. Eid M, Diener E. Handbook of multimethod measurement in psychology. In: Handbook of Multimethod Measurement in Psychology. Washington, DC: American Psychological Association; 2006:141-56.

38. Madera JM, Hebl MR, Martin RC. Gender and letters of recommendation for academia: agentic and communal differences. J Appl Psychol 2009;94:1591-9.

39. Ma A, Koval C, Rosette AS. The Ability-Influence-Means (AIM) model of agency for gender and leadership In submission.

40. Eagly AH, Johannesen-Schmidt MC. The leadership styles of women and men. J Soc Issues 2001;57:781-97.

41. Schaumberg RL, Flynn FJ. Self-reliance: a gender perspective on its relationship to communality and leadership evaluations. Acad Manage J 2017;60:1859-81.

42. Capers Qt, Clinchot D, McDougle L, Greenwald AG. Implicit racial bias in medical school admissions. Acad Med 2017;92:365-9.

43. Sabin J, Nosek BA, Greenwald A, Rivara FP. Physicians' implicit and explicit attitudes about race by MD race, ethnicity, and gender. J Health Care Poor Underserved 2009;20:896-913.

44. Nosek BA, Banaji MR, Greenwald AG. Harvesting implicit group attitudes and beliefs from a demonstration web site. Group Dyn 2002;6:101-15.

45. National Residency Match Program. Results and Data 2016 Main Residency Match. Available at: http://www.nrmp.org/wp-content/uploads/2016/04/Main-Match-Results-and-Data-2016.pdf. Accessed April 12, 2019.