Influence of Gender on Surgical Residency Applicants' Recommendation Letters

Florence E Turrentine, PhD, RN, Caitlin N Dreisbach, MSDS, RN, Amanda R St Ivany, PhD, RN, John B Hanks, MD, FACS, Anneke T Schroen, MD, MPH, FACS

BACKGROUND: Implicit bias has been documented in candidate selection within academic medicine. Gender bias is exposed when writers systematically use different language to describe attributes of male and female applicants. This study examined the presence of gender bias in recommendation letters for surgical residency candidates.

STUDY DESIGN: Recommendation letters for 2016 to 2017 surgery resident applicants selected for interview at an academic institution were analyzed using qualitative text analysis, quantitative text mining, and topic modeling. Dedoose, QDA Miner, and RStudio analytic software were used for analysis.

RESULTS: There were 332 letters of recommendation for 89 applicants (51% male) analyzed. Of 265 letter writers, 86% were male, 21% chairs, and 50% professors. Average word count was 404. Letter writers for male compared with female applicants had a significantly higher average word count (male = 421, SD 144; female = 388, SD 140, p = 0.035). Standout adjectives (eg *exceptional*), reference to awards, achievement, ability, hardship, leadership, scholarship, and use of applicant's name were most often applied to male applicants. Comments on positive general terms (eg *delightful*), grindstone words (eg *hard-working*), physical description, doubt raisers, and work ethic were most often applied to female applicants. Topic modeling and term frequencies revealed achievement words (*performance*, *career*, *leadership*, and *knowledge*) used more often with male applicants, while caring words (*care*, *time*, *patients*, and *support*) were used more often with female applicants.

CONCLUSIONS: Gendered differences examined through language and text exist in surgical residents' recommendation letters. Implementing tools to help faculty write recommendation letters with meaningful content and editing letters for reflections of stereotypes may improve the resident selection process by reducing bias. (J Am Coll Surg 2019;228:356−367. © 2019 by the American College of Surgeons. Published by Elsevier Inc. All rights reserved.)

Letters of recommendation are highly influential in the residency match process, ranking second behind United States Medical Licensing Examination (USMLE) scores as a factor used by surgery program directors to select applicants for residency interviews.[1] This subjective description of the applicant's abilities and character rounds out more objective information provided by a medical school transcript and USMLE Step 1 and Step 2 scores. Unconscious biases may appear in these subjective assessments. Differences in language and length have been identified by gender in letters of recommendation written for faculty applicants in scientific fields, including medicine.[2-4] Few studies have examined the role of gender in letters of recommendation for residency or fellowship applicants in surgical subspecialties,[5-7] but to our knowledge, none yet specifically in general surgery.

Women now comprise nearly 50% of medical students and residents in select specialties. Despite significant changes in the last 25 years, surgery still lags behind,

---

Abbreviations and Acronyms

ERAS   = Electronic Residency Application Service
LDA   = latent Dirichlet allocation
TF-IDF   = term frequency-inverse document frequency
USMLE   = United States Medical Licensing Examination

---

with women constituting 38% of residents and 17% of faculty.[8-10] The changing demographic of surgical residents has been affected by many factors over the last decades, including introduction of the 80-hour work week and accompanying efforts to balance work-life stresses during training.[11] At current rates of change, women are expected to be 50% of surgery residents in about a decade.[9] Advancement to higher ranks or leadership positions in academic medicine, however, has come at a much slower rate for women. Therefore, we sought to look at recommendation letters at the beginning of a surgical career. We hypothesized that differences in language, such as the use of fewer superlatives and fewer references to achievements, and shorter overall length similar to those found in letters for a later career stage would be present in female as compared with male surgery residency letters of recommendation.

METHODS

We analyzed letters of recommendation entered into the Electronic Residency Application Service (ERAS) for 2016 to 2017 applicants selected for surgical residency interview at a single academic institution. Applicant's name, Association of American Medical Colleges number, and ERAS number were redacted from the letters of recommendation. Letters were scanned into the computer. One Note's optical character recognition was used to convert images of portable document format (pdf) files to Microsoft Word documents. Letters were reviewed and corrections were made as needed, with the most frequent errors being "i" recorded as the number l, punctuation changes between periods and commas, phrases before redaction out of sequence, and formatting issues. Corrected documents were imported into Dedoose version 8.0.36, (SocioCultural Research Consultants, LLC, www.Dedoose.com, 2018), a software package that is used for both qualitative and quantitative descriptive analysis. Word count of the recommendation letter was determined with the salutation and signature excluded.

After reviewing the literature, the following codes were selected a priori and entered into Dedoose: ability,[2,12] achievement,[12] agency,[3] applicant's name,[13] clinical skills,[13] communal,[3] doubt raisers,[4] fund of knowledge,[13] grindstone adjectives,[2] judgment,[13] personality traits,[13] research,[2,12] standout adjectives,[2] teaching,[2] technical skills,[13] and work habits.[13] Qualitative codes were refined while reading the letters of recommendation. Category codes that emerged from the data were separately defined and included academic background, authority (of letter writer) established, awards, career choice, community service, family, future promise, hardship, initiative, leadership, legacy (reference to applicant's family member as surgeon or physician), minority or under-represented group, old school, physical description, positive but general terms, presentations, publications, receptive to feedback, recruitment, and scholarship. Variable definitions for these 36 coding categories are detailed in eTable 1.

Letters were reviewed and coded by 2 researchers (FET, ASI). Discussion and review of 10% of cases allowed for consensus in coding. Additionally, a final review of all letters was conducted by FET to ensure coding was applied consistently throughout the dataset.

Data including gender, academic rank, role, and university or hospital of the letter writer were recorded. If gender of the letter writer was not apparent from the name, supplemental information was retrieved from the letter writer's university or hospital website. The academic rank and role of each letter writer was identified from the letter. Rank included professor, associate professor, assistant professor, or no identified rank. Specific roles of the letter writers were noted, including chair, program director, and clerkship director. Applicant demographics, including gender, age, and self-identified race, as well as final rank list quartile after the interview process were uploaded into Dedoose and linked to each letter of recommendation. Rank quartile was designated as Quartile 1: 1 to 22, Quartile 2: 23 to 44, Quartile 3: 45 to 66, and Quartile 4: 67 to 89.

Coded excerpts from each variable above were re-read, and frequently occurring words were entered into QDA Miner Lite v2.0.2. text retrieval software 2018 (Provalis Research, https://provalisresearch.com/products/qualitative-data-analysis-software/freeware/). Each identified phrase was read to ensure appropriate context. Text analysis in QDA Miner Lite was used to explore counts of pronoun use, honorifics, and descriptors contained in letters of recommendation.

In complement, data were read into RStudio (version 1.1.456 at https://www.rstudio.com) to perform text-mining on the letter documents. Textual data and letter metadata, including gender of the applicant and letter writer, were merged in RStudio and further split by gender of the applicant for analysis. The text data were cleaned to remove punctuation, additional spacing, and

Table 1. Surgery Residency Interviewee Demographics

Applicant characteristic	Overall	Female, (n = 44), 49%	Male, (n = 45), 51%
Age, y, mean	26.8	27.3	26.3
Self-identified race, n (%)			
White	58 (65)	32 (55)	26 (45)
Black	7 (8)	5 (71)	2 (29)
Asian	11 (12)	5 (45)	6 (55)
Hispanic	5 (6)	1 (20)	4 (80)
Other	4 (4.5)	1 (25)	3 (75)
Not reported	4 (4.5)	0 (0)	4 (100)

general stop words specific to this corpus, including, but not limited to, *medical, university,* and *student.* Term frequencies were identified using term frequency-inverse document frequency (TF-IDF) to find unique words taking into consideration frequency across all the documents. The TF-IDF measures the importance of a keyword by comparing it to the frequency of the term in a large set of documents. The higher the TF-IDF score, the more rare and distinguishing a word is compared with others. Topic modeling using latent Dirichlet allocation (LDA) was performed on all documents to find generalized themes within and across recommendation letters. Latent Dirichlet allocation is a type of text mining method for developing a topic model, a type of statistical model for discovering the abstract "topics" that occur in a collection of documents. The results produced from LDA are lists of terms that are grouped statistically. The topics are then identified by the individual, not the algorithm. Some topics may seem like incoherent groups of words and others may hint that topics or themes that are defined a priori. The top words within the letters and LDA-identified topics were reviewed.

Descriptive statistics and chi-square test were performed using SPSS Software, (IBM Corp., Released 2017. IBM SPSS Statistics for Windows, Version 25). Letter word count was analyzed by applicant gender, race, and rank list quartile as well as letter writer gender, academic rank, and role. The study was exempted by the University of Virginia Institutional Review Board for the Social and Behavioral Sciences (#2017-0277-00).

RESULTS

For application cycle 2016 to 2017, 89 applicants were selected for interview at our academic surgical residency program, from a total of more than 900 applications. The 332 letters of recommendation submitted in support of these applicants comprised the study dataset. Of these, 329 letters were written by an individual and 3 letters were written by a team of faculty. There were 167 letters of recommendation for 44 female applicants and 165

letters of recommendation for 45 male applicants. The mean number of letters per applicant was 3.7, with a range of 2 to 4. Among these applicants, the average age was 26.8 years, 51% were male, and 65% self-identified race as white (Table 1).

Of 265 letter writers, most were male (228, 86%). Roles held by these male letter writers included 51 chairs, 27 program directors, and 17 clerkship directors. Together, these letter writers with a specific identified role contributed 42% of the recommendation letters. More than half of the male letter writers were professors (122, 54%). The 37 female letters writers included 4 department chairs, 8 program directors, and 3 clerkship directors. Ten of the female letter writers (27%) held the rank of professor.

The letter writers represented 82 institutions, with a range of 1 to 41 and a median of 3 letters per institution. Letters from faculty at the University of Virginia were over-represented in this sample because each of our own student applicants was interviewed. Of the 329 individual letters, most were written by professors (54%), followed by associate professors (22%), assistant professors (15%), and physicians with no specified academic rank (10%). Twenty-three percent of letters were written by a surgery department chair; only 14% were written by a program director, and 8% by a clerkship director. The majority of letters were written by men (290 letters, 88%). There was no difference in percentage of letters written by women for male applicants compared with female applicants (10% vs 13%, respectively, p = 0.39).

Letter of recommendation length

The average word count per letter of recommendation was 404. As measured by average word count, letters written for female applicants were significantly shorter than those for their male counterparts (388, SD 140 vs 421, SD 144, respectively, p = 0.035) (Table 2). Women letter writers tended to use more words, on average, per letter than men (436, SD 159, range 212 to 941 vs 398, SD 139, range 127 to 928, p=0.12). This tendency towards

Table 2. Mean Word Counts Found in Recommendation Letters for Male and Female Surgery Residency Applicants

Variable	Male applicant, no. of letters; (n = 165 letters)	Female applicant, no. of letters; (n = 164 letters)	p Value
Overall	421	388	0.035
Male letter writer	148	142	
Overall	417	377	0.014
Professor	400	369	
Associate professor	409	440	
Assistant professor	526	344	
No specified faculty rank	412	323	
Female letter writer	17	22	
Overall	454	423	0.55
Professor	541	444	
Associate professor	416	413	
Assistant professor	400	428	
No specified faculty rank	446	378	
Quartile, rank list			
1st	468	372	0.0008
2nd	395	407	0.74
3rd	446	370	0.012
4th	384	408	0.4326

a higher average word count was particularly found among letters from women professors compared to men (485, SD 159 vs 386, SD 135, p = 0.02). However, both women and men tended to write longer letters for their male residency candidates than for the female candidates. Men writing letters of recommendations for male applicants compared to female applicants had a significantly higher word count (417, SD 140; 377 SD 135, respectively, p = 0.014). Women writing letters of recommendations for male applicants compared with female applicants also had a higher mean word count, albeit not statistically significant (454, SD = 179; 423, SD 144, respectively, p = 0.55) (Table 2). Letters written by chairs were shorter than those written by others, but this was not statistically significant (mean word count 380 vs 412; p = 0.083)

Three letters were written by teams: 2 teams comprised 1 male and 1 female surgeon, and 1 team consisted of a division of 3 female and 6 male surgeons. Team letter writers produced a mean word count of 624 (range 489 to 761). All team letters were for female applicants.

When considering our final rank list, applicants ranked in the first quartile had the highest mean word count, 418 words, in their recommendation letters. Letters for applicants in the second, third, and fourth quartiles had mean word counts of 401, 402, and 395, respectively. Among male applicants, the candidates ranked in the first quartile also had the highest mean word count in their recommendation letters; the same pattern was not found among the female applicants (Table 2).

Characteristics and language

Certain similarities and differences were found when comparing letters for male and female applicants in terms of characteristics of the applicant's record selected for inclusion in a letter and language used to describe the applicant (Table 3). When featured in a letter, references to academic background, community service, research, presentations, and publications were found in similar proportions between letters for female and male candidates. Additionally, comments referring to knowledge, judgment, technical skills, and teaching were not significantly different in their distribution by gender. Clinical skills were commented on more frequently in letters for male than female applicants, but this did not reach statistical significance. On the other hand, when present, comments referring to ability, achievements, awards, leadership, and scholarship were found significantly more often in letters for male applicants. The use of superlatives and the applicant's name were also more frequently noted in male applicant letters, as was hardship. Grindstone words, positive but general terms, and references to work ethic were featured significantly more frequently in letters for female applicants. Doubt raisers and physical description were also mentioned more frequently in female applicants'

Table 3. Categories of Applicant Characteristics or Language Terms Counted in Surgery Residency Letters of Recommendation by Applicant Gender, Listed by Total Count and Proportion

Category	Total count	Applicant Female, % (count)	Male, % (count)	p Value
Ability	596	46 (277)	54 (319)	0.0058
Top*	133	41 (55)	59 (78)	0.0034
Academic background	172	47 (81)	53 (91)	0.2665
Achievement	139	39 (54)	61 (85)	0.0003
Achieve*	91	41 (37)	59 (54)	0.0155
Agency	236	51 (120)	49 (116)	0.6643
Applicants' name	2,511	46 (1,148)	54 (1,363)	<0.0001
Authority established	111	46 (51)	54 (60)	0.2343
Awards	251	38 (94)	62 (157)	<0.0001
AOA*	52	29 (15)	71 (37)	<0.0001
Career choice	134	48 (64)	52 (70)	0.5134
Clinical skills	279	46 (129)	54 (150)	0.0590
Communal/teamwork	380	53 (202)	47 (178)	0.0983
Community service	127	46 (58)	54 (69)	0.2032
Doubt raiser	99	57 (56)	43 (43)	0.0494
Family	29	45 (13)	55 (16)	0.4503
Fund of knowledge	153	50 (76)	50 (77)	1.0000
Future promise	377	51 (192)	49 (185)	0.5831
Grindstone	337	54 (181)	46 (156)	0.0380
Hardship	21	24 (5)	76 (16)	0.0009
Initiative	153	49 (75)	51 (78)	0.7269
Judgment	8	63 (5)	37 (3)	0.3139
Leadership	170	34 (58)	66 (112)	<0.0001
Leader*	148	41 (60)	59 (88)	0.0020
Legacy	19	58 (11)	42 (8)	0.3304
Minority/under-represented	17	47 (8)	53 (9)	0.7303
Old school	35	49 (17)	51 (18)	0.8681
Personality traits	216	47 (102)	53 (114)	0.2129
Physical description	8	88 (7)	12 (1)	0.0032
Smile*	8	88 (7)	12 (1)	0.0032
Positive, but general	238	56 (134)	44 (104)	0.0089
Delightful*	65	62 (40)	38 (25)	0.0064
Positive*	34	68 (23)	32 (11)	0.0032
Presentations	94	51 (48)	49 (46)	0.7845
Publications	140	49 (68)	51 (72)	0.7383
Receptive to feedback	17	53 (9)	47 (8)	0.7303
Recruitment	188	47 (89)	53 (99)	0.2453
Research	257	51 (132)	49 (125)	0.6506
Scholarship	158	44 (69)	56 (89)	0.0332
Standout adjectives/superlatives	611	40 (242)	60 (369)	<0.0001
Star*	33	33 (11)	67 (22)	0.0061
Impressive*	53	38 (20)	62 (33)	0.0139
Exceptional*	78	37 (29)	63 (49)	0.0012
Excel*	385	56 (214)	44 (171)	0.0009
Bright*	60	62 (37)	38 (23)	0.0088
Teaching/mentoring	68	43 (29)	57 (39)	0.1038
Technical skills	174	51 (88)	49 (86)	0.7095
Work habits/ethic	191	56 (107)	44 (84)	0.0192

*QDA Miner Lite v2.0.2. text retrieval.
AOA, Alpha Omega Alpha honor medical society.

letters. Similar differences by gender were found whether the letter was written by a chair or not (data not shown).

Of note, key characteristics relevant to residency application, namely clinical skills, judgment, technical skills, and knowledge, were not featured in every letter, as marked by counts less than 332. Other than name, characteristics that were found more than 332 in total count included ability, communal/teamwork, future promise, superlatives, and grindstone adjectives. Because each reference was separately counted, these counts over 332 do not indicate that every letter contained reference to each of these categories of characteristics, but these characteristics were commented on with very high frequency.

Leadership was coded in 37% of recommendation letters. Among these 124 letters containing reference to leadership, 60% were for male applicants and 61% were ranked in the first or second quartile. Among the 170 leadership excerpts, male letter writers attributed leadership more frequently to male than female applicants (67% vs 33%, respectively, p < 0.0001). Female letter writers also mentioned leadership characteristics more commonly in letters for male than female applicants (53% vs 32%, respectively, p = 0.1921).

Statements categorized as doubt raisers were coded more frequently in letters for women than men applicants (57% vs 43%, p = 0.0494). Some of these statements were quite candid, such as, "His psychomotor skills are at least at the level of his peers," or "I have actually had to encourage him to be more confident," whereas others included qualifiers. Although the exact intent of the qualifier is difficult to discern, it may serve as a subtle doubt raiser. Examples of this include the following:

"The fact that ….does not yet have a completed, and published, research paper has nothing to do with her lack of commitment…"

"While she perceived herself as disadvantaged compared to her peers going into surgery as she had not practiced technical skills,…"

"…will never shirk responsibility, especially when there is nothing for her to gain personally."

Statements made about male applicants seemed more often declarative, while similar positive statements supporting female applicants seemed more likely to include a qualifier. Examples of this include the following:

"Academically, he is a STAR. He is a rising STAR in surgery and an exceptional young man." "Simply put, she is a budding SUPERSTAR and with the appropriate guidance, I think she will clearly be one of the leaders in surgery in the future."

In closing the letters, writers mentioned intending to recruit the applicant to stay at their institution for residency 188 times. These remarks were found similarly in letters for male and female applicants. Chairs included interest in recruiting 46 applicants, the majority of whom were male (65% vs 35%, p = 0.0042). Providing contact information and inviting further discussion of the applicant was included in letters authored by professors significantly more often than assistant professors (46% vs 18%, respectively, p < 0.0001). The closing comment, "If I can provide additional information…," was found in 97 letters and trended toward being seen more frequently in letters for female applicants (54 of 97 letters, 56% vs 44%, p = 0.0955). Letters closing with this commonly used phrase were more often seen among applicants ranked in the third or fourth quartile, as compared with those in the first or second quartile (58% vs 42%, respectively, p = 0.0262).

Term frequency-inverse document frequency analysis and topic modeling

Term frequency-inverse document frequency was used in this analysis to statistically define the words that are most distinguishing among recommendation letters from each gender represented in the data. A high TF-IDF statistic means that a particular word has a strong relationship within a document or corpus of documents while filtering out words that occur frequently. For male gender applications, TF-IDF analysis isolated high-ranking words such as *best*, *career*, *leadership*, and *skills*. In comparison, female applicant documents showed words such as *care*, *good*, *team*, and *opportunity* (Table 4). Words such as *patients* and *knowledge* were highlighted between both documents. In addition to TF-IDF, topic modeling resulted in loosely defined topics that included groups of words that are mathematically associated with each other. The LDA-grouped words such as *career*, *degree*, *faculty*, and *patients* appeared across the combined male and female dataset. More specifically, female applicant letters showed topics

Table 4. Comparison of Ranked Terms Identified by Term Frequency—Inverse-Document Frequency Analysis by Gender

Male applicant	Female applicant
Patients	Care
Society	Patients
Surgeon	Surgeon
Best	Team
Career	Good
Leadership	Time
Knowledge	Knowledge
Demonstrated	Opportunity
Skills	Questions

related to work ethic, care giving, and support, while male applicant letters had topics such as performance and technical skill.

DISCUSSION

Letters of recommendation written for general surgery residency applicants demonstrate significant similarities and differences in language by gender. References to ability, achievements, awards, leadership, and scholarship were found more commonly in letters about male applicants; references to work ethic were more common in letters about female applicants. Words categorized as superlatives (eg *stellar, unparalleled*) were counted more frequently in letters for male applicants; words categorized as grindstone adjectives (eg *dedicated, conscientious*) or positive but general (eg *delightful, enthusiastic*) were counted more frequently in letters for female applicants. Although all of these words are intended as commendations, these patterns in word selection follow gender schemas of associating achievement and ability with men and effort and nurturing with women.[4,14] Gender schemas, as described by Valian,[14] refer to "a set of implicit, or nonconscious hypotheses about sex differences that play a central role in shaping men's and women's professional lives…affect our expectations of men and women, our evaluations of their work, and their performance as professionals."[14]

The language patterns found in our study are similar to those described in works examining letters at later career stages in other scientific fields.[2-4,15] In particular, superlatives or standout adjectives were more commonly used in letters for males, as was also found by Schmader and colleagues.[2] Describing leadership ability more commonly for men than women is also found in other studies.[5,6] This is particularly noteworthy given the slow advancement of women to higher leadership positions in surgery.[9] Furthermore, letters for male applicants were longer and included more references to the candidate by name. Although our study was not designed to identify the impact of recommendation letters in receiving an invitation to interview, differences such as those identified in this study are recognized to be important. In one of few studies focused specifically on general surgery residency letters of recommendation by Greenburg, those letters rated highest were twice as long as lower rated letters, contained more personal references by name, and featured more superlatives or comparisons to a reference point indicating the candidate to be stellar.[13]

In our comparison of human coding and computer text mining methodologies, consistency was demonstrated in leadership being associated with letters for men and work ethic with letters for women. A thematic difference attributing performance with men and caring with women also was identified. Our comparisons of these qualitative and quantitative methods show that they can corroborate one another. However, the detail and richness of content detailed in the manual coding could not be replicated with text mining and topic modeling. In future research, the topics identified as important in the qualitative analysis could be used to guide interpretation of the computer text mining findings, allowing a far more efficient methodology to be used in place of the time-intensive qualitative coding and analysis.

Recommendation letters were inconsistent in providing commentary on key applicant characteristics relevant to surgery residency application, such as clinical skills, judgment, technical skills, and knowledge. Instead, letters often focused on generalities, as marked by frequently using words categorized as grindstone adjectives or superlatives, or words related to future promise and teamwork. Focusing on personality traits and generalities may indicate not knowing the applicant very well. Shortened surgical clerkships and a tendency to ask more senior faculty for recommendation letters in hopes of greater national name recognition may both play into the letter writer not knowing the applicant particularly well.

Our study identified some important similarities between letters for male and female applicants. Differing from previous studies, there was no difference by gender in frequency of references to application of agency or communal/teamwork skills.[3,16] This may reflect the value placed on a drive to perform and functioning well within a team as characteristics for an ideal surgery resident.[17] Other noteworthy similarities include comparable references to research, technical skills, knowledge, future promise, publications, initiative, and teaching. In evaluating letters for chemistry and biochemistry faculty applicants, Schmader and colleagues[2] also found significant similarities between men and women applicants.

Our study showed statements indicating intent to recruit the applicant to stay at the writer's institution for residency appear in similar frequency in recommendation letters for men and women. Recruitment is viewed as one of the most positive elements to be found in a surgical residency letter of recommendation.[13] Closing the letter with an offer to provide additional information if any questions was found more commonly in letters for women applicants. Interpretation of this is unclear. This closing was viewed as having negative connotations in the study by Greenburg and colleagues,[13] but is suggested as a routine closure in several recommendation letter templates.[18,19]

As with some previous studies,[4,13] we found letters for male applicants to be generally longer than those for females. Longer letters have been associated with higher favorability and greater caring about the recommendee.[4,13,20] Shorter letters tend to be more general, and focus on personality traits, with less detail describing applicant's knowledge or skills.[13] Brevity and generality are clearly seen as reflecting negatively on the applicant. Regardless of length, a meaningful letter of recommendation is suggested to have the following 3 components: the commitment and relationship of the letter writer to the applicant, specific comment reflecting the record of the applicant, and a critical evaluation or comparison of the qualities and achievements of the applicant.[4]

The following examples from our study highlight other considerations for improving the quality of recommendation letters for residency. Generic terms about personality (eg "personable and well-liked") and work habits (eg "very hard working") could be replaced with substantive summations of the candidate's work (eg "She is committed to her volunteer work, especially providing CPR lessons and serving as a mentor to students interested in the medical field."). Similarly, grindstone adjectives (eg "dependable," "organized") typically describe very desirable traits in a prospective resident, but are overused and thereby less meaningful unless these words are contextualized. For instance, "…engaged and effective in her work" is less informative than "…clinical skills were clearly advanced; in fact I found his daily patient notes more accurate, more focused and more helpful than those of some of our residents," which provides specific context, an evaluation, and a comparison. Qualifying statements (eg "…with the appropriate guidance, I think she will clearly be one of the leaders in surgery in the future"), although true, can serve as doubt raisers and should be used with caution. References to personal life or physical attributes were uncommon in our study but stood out when present (eg "…she is married and has a husband with a strong support of her surgical career."; "…slender, attractive, rather tall individual with sandy brown hair and a smile that radiates in the room.") and are probably best avoided.

Clearly, writing a well thought-out, honest, and genuine letter requires knowing the candidate well. To this end, we suggest that programs requiring a letter from the chair reconsider this practice, instead allowing the applicant to select the 3 faculty members who can provide the most insightful evaluations for their letters. Greenburg and colleagues[13] found that although faculty reviewers in their study stated upfront that academic rank of the letter writer was important, when provided full and deidentified letters to score on quality, reviewers did not score these letters differently, suggesting that rank of the writer is not as important as believed.[13] Furthermore, applicants are encouraged to seek out faculty with a reputation for writing excellent letters and to prepare readily usable depictions of their accomplishments or strengths. Letter writers need to assure that the basic elements of the letter, including a genuine assessment of the applicant's knowledge, judgment, and clinical skills, as well as acknowledgement of the applicant's most notable past accomplishments or character traits, are present. An excellent example of this reads, "I would like to focus on her attributes as they demonstrate the qualities necessary for a surgical residency: outstanding knowledge, clinical skills, patient care, leadership, and interpersonal skills." Omissions of key elements result in a letter of minimal assurance and leave the reader paying more attention to what is not said as opposed to what is written.[4] The letter should be carefully proofread; errors, such as in spelling or pronoun mismatches, may still reflect negatively on the applicant rather than the writer. Judicious and honest use of key phrases, such as a wish to recruit the applicant to one's own institution, rating the applicant as one of the best, or comparing performance to that of a resident, can heavily influence the impact of the letter.

Several tools exist to help reduce evidence of gender schema in letter content and word selection. Awareness that differences in language by gender have been found in numerous settings is essential to recognizing and reducing these biases.[2-6,15] We are especially prone to fall back on gender schemas in writing recommendation letters when pressed for time and when we do not know the applicant particularly well.[4] Educational resources, such as a 20-minute educational intervention, have been shown to successfully change faculty members' perceptions of bias.[21] Furthermore, general implicit bias training is recommended. Standardized letters of recommendation[22-25] or having letter writers address Accreditation Council for Graduate Medical Education competencies in letters may also reduce bias.[26] Some specialties, such as otolaryngology, have adopted a standardized letter of recommendation. Evaluation of the standardized, as compared to the narrative portion of the recommendation letters, for otolaryngology residency demonstrated a reduction in gender biases reflected in the standardized letters, but certain differences in how male and female applicants were described remained.[5] Nevertheless, the authors of this study concluded that retaining the narrative portion of the recommendation letters was essential to conserving the value of the recommendation letter in

the application.[5] Efforts should be directed at helping letter writers recognize how unconscious biases can be reflected in language and content selection. Providing a tip sheet that highlights common issues may be beneficial.[27] Using an on-line bias calculator to review letters of recommendation could also alert one to language use that is skewed toward words that are more commonly associated with men or women and provide an opportunity to reevaluate the letter before it is sent.[28]

Limitations of our study include that our findings reflect letters received in a single year, by a single institution, and represent only applicants selected for an interview. Because our residency program interviews all of our own medical students applying to general surgery residency, letters from our own faculty were heavily represented in this sample. Analysis was limited to the content of the recommendation letters, and supporting data from curriculum vitae were not reviewed. However, this represented a highly select and competitive group of applicants because only those selected for interviews were included. Pronouns were not removed from letters so coder bias is a consideration. However, we elected to conduct this study with coders who are not surgeons, thereby reducing bias that could arise when evaluating one's own profession or peer group. We used the first race term that applicants selected to identify themselves and did not account for multiple race terms. Race and ethnicity of the applicants or those who wrote letters of recommendation were not analyzed because our study sample was too limited to permit this analysis.

CONCLUSIONS

Letters of recommendation written for general surgery residency applicants demonstrate significant similarities and differences in language by gender. Important differences include greater prevalence of references to leadership and achievement for men and work ethic for women, include more frequent use of superlatives for men and grindstone adjectives and positive but general terms for women, and longer letters for men. These differences are similar to those found in recommendation letters for later career stages. Recommendation letters were inconsistent in providing commentary on key applicant characteristics relevant to surgery residency application, such as clinical skills, judgment, technical skills, and knowledge. Instead, letters often focused on generalities, likely reflecting not knowing the applicant particularly well. Important similarities between letters for male and female applicants included comparable references to teamwork skills, research, knowledge, future promise, and technical skills. Substantive content in

recommendation letters could be improved by avoiding generic terms about personality and work habits, focusing on specific accomplishments or work, using grindstone adjectives within context, providing summative evaluations and comparisons, and avoiding unnecessary qualifying statements. Implementing tools to help faculty write recommendation letters with meaningful content, and edit letters for reflections of gender schema, may improve the resident selection process and reduce the impact of bias at the start of a surgical career.

Author Contributions
Study conception and design: Turrentine, Dreisbach, St Ivany, Schroen
Acquisition of data: Turrentine
Analysis and interpretation of data: Turrentine, Dreisbach, St Ivany, Hanks, Schroen
Drafting of manuscript: Turrentine
Critical revision: Dreisbach, St Ivany, Hanks, Schroen

REFERENCES
1. National Resident Matching 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 October 30, 2018.
2. Schmader T, Whitehead J, Wysocki VH. A linguistic comparison of letters of recommendation for and female chemistry and biochemistry job applicants. Sex Roles 2007;57:509−514.
3. Madera JM, Hebl MR, Martin RC. Gender and letters of recommendation for academia: Agentic and communal differences. J Appl Psychol 2009;94:1591−1599.
4. Trix F, Psenka C. Exploring the color of glass: Letters of recommendation for female and male medical faculty. Discourse Society 2003;14:191−220.
5. Friedman R, Fang CH, Hasbun J, et al. Use of standardized letters of recommendation for otolaryngology head and neck surgery residency and the impact of gender. Laryngoscope 2017;127:2738−2745.
6. Hoffman AL, Grant WJ, McCormick MF, et al. Gendered differences in letters of recommendation for transplant surgery fellowship applicants. Academic Surgical Congress Abstracts Archive. 2018. Available at: http://www.asc-abstracts.org/abs2018/77-03-gendered-differences-in-letters-of-recommendation-for-transplant-surgery-fellowship-applicants/. Accessed January 16, 2019.
7. 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;118:1335−1344.
8. Sexton KW, Hocking KM, Wise E, et al. Women in academic surgery: The pipeline is busted. J Surg Educ 2012;69:84−90.

9. Abelson JS, Chartrand G, Moo T, et al. The climb to break the glass ceiling in surgery: Trends in women progressing from medical school to surgical training and academic leadership from 1994 to 2015. Am J Surg 2016;212:566−572.

10. Davis EC, Risucci DA, Blair PG, Sachdeva AK. Women in surgery residency programs: Evolving trends from a national perspective. J Am Coll Surg 2011;212:320−326.

11. Wasserman M. Hours constraints, occupational choice, and gender: Evidence from medical residents. Available at: http://conference.iza.org/conference_files/Gender_2018/wasserman_m26072.pdf. Accessed September 14, 2018.

12. Issac 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.

13. Greenburg AG, Doyle J, McClure DK. Letters of recommendation for surgical residencies: What they say and what they mean. J Surg Res 1994;56:192−198.

14. Valian V. Why So Slow? The Advancement of Women. Cambridge, MA: MIT Press; 1997.

15. Dutt K, Pfaff DL, Bernstein AF, et al. Gender differences in recommendation letters for postdoctoral fellowships in geoscience. Nature Geoscience 2016;9:805−808.

16. Phelan JE, Moss-Racusin CA, Rudman LA. Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. Psychol Women Quarterly 2008;32:406−413.

17. Lv J, Guerlain S, Turrentine B. Do personality and level of experience affect perceived safety culture? A study of surgical residency programs across Virginia. Safety Across High-Consequence Industries 2011;1:22−39.

18. Useful phrases for recommendation letters. eduers.com Web site. Available at: https://www.eduers.com/graduate/useful_phrases_recommendations/. Accessed September 14, 2018.

19. Writing letter of recommendation. Sample letter of recommendation for residency program. MITE Web site. Available at: https://www.mitemmc.org/resources/educators-tool-box/writing-letter-recommendation/. Accessed September 14, 2018.

20. Judge TA, Higgins CA. Affective disposition and the letter of reference. Organizational Behavior Human Decision Processes 1998;75:207−221.

21. Girod S, Fassiotto M, Grewal D, et al. Reducing implicit gender leadership bias in academic medicine with an educational intervention. Acad Med 2016;91:1143−1150.

22. Keim SM, Rein JA, Chisholm C, et al. A standardized letter of recommendation for residency application. Acad Emerg Med 1999;6:1141−1146.

23. Alweis R, Coltichio F, Milne CK, et al. Guidelines for a standardized fellowship letter of recommendation. Am J Med 2017;130:606−611.

24. Love JN, Ronan-Bentle SE, Lane DR, Hegarty CB. The standardized letter of evaluation for postgraduate training: A concept whose time has come? Acad Med 2016;91:1480−1482.

25. American College of Surgeons. (2018). Standardized reference letter for medical students applying to general surgery programs. Available at: https://www.facs.org/∼/media/files/about%20acs/governors/5surgerystandletrec.ashx. Accessed December 4, 2018.

26. Stohl HE, Hueppchen NA, Bienstock JL. The utility of letters of recommendation in predicting resident success: Can the ACGME competencies help? J Grad Med Educ 2011;3:387−390.

27. Commission on the Status of Women. Avoiding gender bias in reference writing. Available at: https://csw.arizona.edu/sites/default/files/avoiding_gender_bias_in_letter_of_reference_writing.pdf. Accessed September 12, 2018.

28. Forth T. Gender bias calculator. Gender Bias Calculator Web site. Available at: https://www.tomforth.co.uk/genderbias/. Accessed September 12, 2018.

Discussion

DR GEORGE SAROSI, JR (Gainesville, FL): Dr Schroen and colleagues have presented an interesting content analysis project looking at sex differences in letters of recommendation for medical students applying for surgical residency. Notable findings of the work are that letters for male applicants are longer than for female applicants and are focused on the applicant's achievements, while letters for female applicants are focused on personality traits such as determination. What is interesting is that both male and female letter writers display the same behaviors or biases.

You have used position on your rank list as a measure of applicant achievement. Although you have equal numbers of men and women in your total sample, are they evenly distributed across the rank quartiles? Do you have data on the sex distribution of each rank quartile? Do women or men get ranked higher? My bias would be that women get ranked higher, which makes your sex differences even more striking based on my experience as a program director. Did you consider collecting some objective data on applicant achievements such as United States Medical Licensing Examination (USMLE) scores, class rank, or number of publications? Is there a difference between the sexes? And if so, is this reflected in the letters? Does your rank list also reflect this? One quarter of your letters are written by chairs of surgery. Presumably, they write more letters than other faculty. As a subgroup, are the letters from chairs different than the other letters written by faculty? Is there some influence of experience on letter writers?

Finally, based on your findings that comments about the important traits of applicants such as clinical skill, knowledge, judgment, and technical skills are not found in all or even most letters, do you have recommendations for guidelines that all letter writers should use when preparing surgery letters of recommendation? More importantly, should surgical organizations consider publishing such guidelines to try and combat these biases?

DR JULIE ANN SOSA (San Francisco, CA): This research confirms the findings of others, this time in surgical education, that there is a measurable appearance of implicit sex bias in letters of recommendation for general surgery residency candidates as demonstrated by 2 things: male candidates having longer letters written for them by both male and female faculty members; and more standout adjectives being applied to male candidates and more general and grindstone adjectives used to describe women.

eTable 1. Variable Categories, Definitions, and Study Examples

Category	Definition	Exemplar
Ability[2,12]	Talent, skill, or proficiency in an area. Adept, competent, skilled, ability, bright, able, analytical, smart, proficient, capable, able, talented, intelligent, genius, brilliant, brain, aptitude, gift, capacity, propensity, innate, flair, knack, clever, expert, proficient, natural, inherent, instinct, adroit, creative, insight	"brilliant young man"; "very intelligent individual"; "extremely competent"
Academic background	Formal preparation before medical school including graduate and undergraduate degrees	"…great foundation to enter surgery residency. After attending magnet high school, she attended *University where she graduated cum laude."; "graduate of the University of *School of Business"
Achievement[12]; athletics; mastery	Successful effort. Organized, performed, earned; varsity or professional sport; professional/national competition	"currently second out of 121 in his class"; "varsity tennis player"; "number 1 student at University of *"
Agency[3]	Drive to perform, achieve status; confident, ambitious, competitive, assertive, independent, decisive, aggressive, dominant, forceful, daring, outspoken, intellectual	"…passion, stamina and drive he has, as he has truly worked his way from the ground up."; "possesses a tenacity which will serve her well throughout her career"
Applicants' name[13]	Use of applicant's name	"I give — my highest recommendation…"
Authority established	Letter writer outlines their qualifications: in my 20-year career, my decades of experience	"Presently, I am a practicing general surgeon in …, formerly the Chairman of the Department of Surgery and President of the Medical Staff at ….hospital. I am an associate clinical professor of general surgery at…"
Awards	Mark of recognition; scholarships, Alpha Omega Alpha Honor Medical Society membership	"…awarded the prestigious *Medical Student Scholarship Award for her superior research and academic efforts plus her dedication to surgical care of injured patients."; "earned the Gold Humanism Honor Society and the Commendation for Outstanding Academic Performance"
Career choice	Surgical subspecialty, academic surgery	"Plans to pursue a career in academic surgery and I believe he has all the tools to succeed."; "additional training after general surgery, most likely in cardiothoracic surgery"
Clinical skills[13]	Observable act within patient care.; physical exam, diagnostic, communication, rounding	"…clinical skills were clearly advanced; in fact I found his daily patient notes more accurate, more focused and more helpful than those of some of our residents."; "engaged and effective in her work on the wards"
Communal/ team work[3]	Working together; team work or team player; modest, cooperative, competent, helpful, kind, sympathetic, sensitive, nurturing, affectionate, agreeable, tactful, interpersonal, warm, caring	"consummate team player"; "great team player and is always asking the residents how she can help out"
Community service	Volunteer work to help others	"She is committed to her volunteer work, especially providing CPR lessons and serving as a mentor to students interested in the medical field."; "volunteering at his place of worship and at one of the city's kitchens for the homeless"
Doubt raiser[4]	Negative language, hesitancy, hedges, faint praise, irrelevancies; not complementary or supportive of candidate.	"I have actually had to encourage him to be more confident…"; "has one Achilles heel on his application. His part I score is low."; "limitations evidently perhaps keeping him from matching were issues of maturity and communication"

(Continued)

eTable 1.   Continued

Category	Definition	Exemplar
Family	Mention of spouse, children, parents	"….married to a * in town and they currently have no children."; "she is married and has a husband with a strong support of her surgical career"
Fund of knowledge[13]	Information; clinical knowledge	"… demonstrates an excellent knowledge base."; "excellent fund of medical knowledge"; "superior knowledge"
Future promise	Future potential as surgeon	"… she will be an amazing surgeon."; "will be a fine surgeon, an excellent resident and a contributor to the surgical community for years to come"
Grindstone[2,12]	Hard working, conscientious, dependable, meticulous, thorough, diligent, dedicated, careful, reliable, effort, assiduous, trust, responsible, methodical, industrious, busy, work, persist, organized, disciplined	"very hard working"; "very dependable"; extremely well organized"; worked very, very hard"
Hardship	Difficult time or adversity applicant faced	"Twelve years ago he arrived in * not knowing any English, and with few resources."; "had to overcome much adversity throughout his life"; "lived through Hurricane Katrina and persevered"
Initiative	Assess and undertake independently; task, research, project conceived/enacted by applicant	"…he came to me in his fourth year and asked if he could have teaching sessions with the surgical clerkship students."; "she developed an internship for undergraduate students interested in a medical career"
Judgment[13]	Considered decisions	"Her judgment is excellent."; "displayed knowledge and judgment consistent with a good * medical student"
Leadership	Leading or guiding a group.; president/officer of organization	"…demonstrated exemplary leadership skills."; "As a 4th year student leader she was an OSCE Skills Session Leader"
Legacy	Applicants' parents are physicians/surgeons	"… the daughter of a general surgeon, she has a good idea of what the life entails."; "He comes from a lineage of highly accomplished surgeons"
Minority or under-represented group	Racial or ethnic minority, socioeconomic disadvantage	"…served on the medical school Admissions Committee, as well as leading organizations of Native American diversity, championing her background."; "born in India, he immigrated to the US"; "speaking very little English, boarded a plane and left his mother and homeland"
Old school	Traditional, the way we were trained, back in the day	"She was tireless, often the first to arrive and the last to leave."; "always staying to get the work done, always willing to go the extra mile, and never complaining"
Personality traits[13]	Patterns of thoughts, feelings, behaviors- introversion, openness, conscientiousness, agreeableness	"…affable and very enthusiastic."; "general quiet nature"; "personable and well liked"; "quick with humor"
Physical description	Defining traits of the applicant's body, ie hair color, height, weight, smile	"…slender, attractive, rather tall individual with sandy brown hair and a smile that radiates in the room."; "dutifully went about her team business with a glowing smile"
Positive but general terms	General affirmations; good, excellent, eager, enthusiastic, happy, delightful	"…consistently enthusiastic and pleasant…"; "positive force"; "full of energy
Presentations	Presenting a topic locally or nationally	"Presented her research at the * earlier this year."; " first author on abstract presented at the *"
Publications	Journal articles, manuscripts	"He has 4 first-author publications (of 7 total) published papers."; "She has contributed 25 publications to the peer-reviewed literature!"

(Continued)

eTable 1. Continued

Category	Definition	Exemplar
Receptive to feedback	Open to critique, criticism, feedback	"…always open to constructive criticism."; "willing to improve on hers strengths and accept feedback to improve her weaknesses"
Recruitment	Enlist or retain applicant at home institution	"We will recruit her to stay in our program."; "will be heavily recruited from our institution for residency position"; "we will try to entice him to stay in our residency program"
Research[2,12]	Studious inquiry; study, studies, project, research, fund, science, data, experiment, test, result, finding, publication, publish, vita, method, grant, manuscript, journal, theory, discovery, contribution, science	"…she has performed translational research in the diabetes lab…"; "participated in the * Department of Surgery Medical Student Summer Research Program"; "worked on…complex analysis of the tumor microbiome"
Scholarship	GPA, STEP scores, Honors	"…has USLME scores of 229 on Step I and 255 on Step II. His medical school GPA is 3.9."; "Scored 100% on her oral exams"
Standout adjectives[2,12]/ superlatives	Notable, remarkably superior; excellent, superb, outstanding, unique, exceptional, unparalleled, most, wonderful, terrific, fabulous, magnificent, remarkable, extraordinary, amazing, supreme	"phenomenal"; "exceptional"; "superb"; "fantastic"; "outstanding"
Teaching[2,12]/ mentoring	Teaching, education, training, instruct, supervise, lecture, mentor, adviser, class, counselor, syllabus, course, colleague, citizen, student,	"… active in teaching students surgical techniques via formal knot-tying and suturing classes."; "head teaching assistant for our first year medical students"
Technical skills[13]	Surgical skills, spatial, hand eye coordination, dexterity	"…very capable and technically skillful in everything we allowed her to do- including sternotomy, skin closure, etc."; " "a natural" as far as technical skills are concerned"; "has precise hands and good instincts in the operating room"
Work habits[13]/ ethics	Behavioral, ethical, and practical elements applied by employees in contributing to job performance; sets a precedence of efficiency, productivity and reliability.	"…solid work ethic and is ever present in the OR, in the ICU, and on the floors."; "spent more hours at the hospital, scrubbed in on more cases… than any student"