What's in a Word? Qualitative and Quantitative Analysis of Leadership Language in Anesthesiology Resident Feedback

Nicole Arkin, MD
Cara Lai, BS
Larissa Miyachi Kiwakyou, MD
Gregory Milo Lochbaum, MD
Audrey Shafer, MD
Steven K. Howard, MD
Edward R. Mariano, MD, MAS
Magali Fassiotto, PhD

ABSTRACT

Background Individuals who have agentic traits (eg, assertive, confident, competent) that are more commonly associated with men are often selected for leadership roles. For women, this poses a potential barrier to entry into the higher ranks of academic medicine.

Objective We analyzed anesthesiology resident feedback for differences in the use of agentic descriptors using qualitative and quantitative methods based on resident gender and year of training.

Methods This study uses textual analysis of 435 assessments of residents over a 1-year period within a single residency program. We performed a qualitative content analysis on the words used in resident feedback and performed negative binomial regression analyses to determine significant differences in the way residents were described based on gender and year of training.

Results Female residents were less likely than male residents to be described as agentic after controlling for excerpt length, year of training, and evaluator variability ($\beta = -0.347$; 95% confidence interval [CI] −0.666, −0.028; $P = .033$). Senior residents were more likely to be described as agentic ($\beta = 0.702$; 95% CI 0.402–1.002; $P < .001$) compared to junior residents. The increased number of agentic codes among senior residents was driven by increased agentic description of female residents' ratings in the senior cohort ($\beta = 0.704$; 95% CI 0.084–1.324; $P = .026$).

Conclusions Female residents were described as agentic less often than male residents in early years of training, but the gap was not present among senior residents.

Introduction

The representation of women in medicine has changed significantly in recent decades, with 46.3% of female medical school graduates in 2015–2016 compared with 6.9% in 1965–1966. However, women still comprise only 22% of full professors in academic medicine.[1,2] The field of anesthesiology has not seen an increase in the number of female full professors between 1985 and 2006, despite an increased number of female anesthesiology residents.[1,2] More men than women continue to choose a career in academic medicine, maintaining the gender gap at a time when women's voices in leadership are needed more than ever.[3]

A robust body of research from sociology, management, academics, and now medicine reveals that men and women are perceived differently when it comes to the assessment of leadership traits.[4–7] Leadership roles are generally given to individuals with agentic qualities, such as independence, assertiveness, competence, and efficacy.[8,9] Traditionally, such traits have been stereotypically associated with masculinity, while their communal counterparts of empathy, friendliness, and cooperativity have been stereotypically associated with femininity.[4] Women are less likely to be perceived as portraying these agentic leadership traits; when they do, they can be penalized for going against expectation in a phenomenon known as "role incongruity."[10] By choosing certain words to describe individuals—and omitting others—we may reinforce cultural constructs that become relevant when examining possible barriers to leadership for women in academic medicine.

This study aims to analyze the content of resident feedback using an established construct for agentic leadership from sociological and management literature.[9,11,12] We analyzed the language attending anesthesiologists used in written feedback to residents during clinical anesthesiology training. We hypothesized that feedback given to women would contain fewer agentic descriptors than feedback given to men (consistent with current leadership literature). This difference also would not be present among senior

DOI: http://dx.doi.org/10.4300/JGME-D-18-00377.1

Editor's Note: The online version of this article contains an example of the coding scheme.

residents who are familiar to faculty, so their assessments of performance would be less influenced by gender stereotypes.

What was known and gap
Leadership roles are often given to individuals who have agentic traits that are commonly associated with men, and assessments of residents might be influenced by gender stereotypes, making it harder for women to enter higher ranks of academic medicine.

What is new
An analysis of feedback given to residents for differences in the use of agentic descriptors based on resident gender and year of training.

Limitations
Single site, single specialty limits generalizability, and lack of longitudinal data limits conclusions about training year.

Bottom line
Female residents are less likely to be described using agentic language during the first 2 years of training, but this gender gap is not found among senior residents.

Methods

Setting and Participants

This study used textual analysis of attending anesthesiologists' assessments of anesthesiology residents at a suburban academic teaching hospital over a 1-year period (2016–2017). All clinical anesthesiology residency levels beyond internship were represented: postgraduate year 2 (PGY-2) to PGY-4. No PGY-1 residents rotated in the anesthesiology department at this site.

Study Design

Qualitative content analysis was performed on the types of words used in resident feedback. Resident feedback was analyzed using an "agentic" versus "communal" framework well described in sociological literature and increasingly common in medical literature.[11–13] Code counts were used for quantitative statistical analyses, examining whether there were significant differences in the descriptions of male versus female and junior versus senior residents. The feedback represented both junior and senior training levels to account for differences in the standard against which residents may be evaluated.

Data Collection

All attending anesthesiologists' resident assessments made from January 2016 to January 2017 were obtained from a Veterans Affairs (VA) teaching hospital with an anesthesiology residency program accredited by the Accreditation Council for Graduate Medical Education. Each evaluation represented the assessment of an individual resident at the end of a scheduled anesthesiology rotation by an individual attending anesthesiologist. Evaluators submitted free-text feedback to the rotation director for review and were not provided a standard format. No feedback was excluded from qualitative analysis, but assessments by attending physicians who did not complete more than one were excluded from quantitative analysis.

Coding

The qualitative coding team included 1 anesthesiology resident (N.A.) and 2 medical students (C.L. and L.M.K.) who coded assessments after collection and deidentification. Evaluators were blinded to resident gender, training year, and evaluator gender. A doctoral-level expert in sociological research methods (M.F.) advised the team at all stages. All assessments were coded using the qualitative analysis program Dedoose version 7.0.23 (SocioCultural Research Consultants, Hermosa Beach, CA).

Prior to coding, words typically classified as agentic or communal in sociology and management literature were identified to create a preliminary coding scheme. Using the preliminary scheme, all team members independently coded the first 100 assessments to identify additional themes, establish consensus about modifications to the codebook, and add new codes. During this review process, the thematic category of "technical" was established, defined by anesthesiology-specific feedback regarding residents' technical skills.

The coders discussed at length the specific language in the assessments regarding communication. For an anesthesiology resident, learning to communicate with surgeons requires confidence, competence, and a strong sense of self. Similarly, a critical aspect of anesthesiology is planning. Clinically competent residents are expected to be able to develop a plan for each anesthetic. Effectively communicating the plan demonstrates agentic traits such as competence and leadership in the operating room. The coders made a decision by consensus to include "communicating well with the surgical team" and "effectively communicates plan" because they were determined to require application of agentic traits.

Two subcategories were added to address feedback describing *personality* traits versus observed *behaviors*, a distinction in sociological literature that refers to inherent personality traits versus malleable behaviors.[14,15] Additionally, each excerpt was assigned the descriptor "vague" or "specific" to capture the level of detail and actionability of feedback. Vague feedback included comments, such as "good job," "on top of everything," and "on track," without additional

explanation as to the context of the comment. Specific feedback included "anesthetic plan appropriately adjusted to specific patient issues," "seemed to engage in proactive communication with surgical team and operating room staff," and "responded well to changes in the operating room course, seemed mentally prepared for such."

Following finalization of the codebook, in line with accepted qualitative research methods, interrater reliability was calculated among all members of the coding team using excerpts from the first 49 assessments, resulting in a pooled κ score of > 0.80, which is consistent with accepted standards.[16–18] Discrepancies in coding were discussed, with no further modification to the coding scheme (TABLE 1). The final coding scheme was used to independently code the remaining 386 evaluations, divided evenly among coders. A second interrater reliability test was applied to 10 random samples from each coder's set of evaluations (a total of 30), which again resulted in a pooled κ of > 0.80 for all raters.

Thematic analysis was performed on assessment excerpts, with each excerpt defined as a word or phrase with a unique thematic element. A single assessment could contain multiple excerpts, and each excerpt was assigned 1 code from every column in the coding scheme (provided as online supplemental material). The code counts were then totaled for each assessment. The fully coded data set was then unmasked to resident gender, level of training, and linked to the individual attending to create the final data set prior to analyses.

The Stanford University Institutional Review Board and VA Palo Alto Research Committee approved this study.

Statistical Analysis

We combined PGY-2 and PGY-3 residents into a junior cohort to be compared with PGY-4 residents in the senior cohort. Roles differ for junior versus senior residents during rotations at this site. The PGY-2 and PGY-3 roles at this site are similar, as the residents perform anesthetics in a comparable range of surgical cases. PGY-4 residents often elect into this rotation, are assigned the most medically and technically complex cases, and take on a more supervisory role, providing lectures and facilitating learning for the junior residents. Therefore, we hypothesized that PGY-4 evaluations (compared with those of junior residents) might have different expectations reflected in their assessments.

We developed a multivariate regression model to evaluate the effect of residency training and gender on resident feedback. Dependent variables were counts for each category (agentic, communal, technical, other, personality, behavior, specific, vague) within a given assessment. Since the dependent variable was a count variable limited in range, a negative binomial regression model was used. In all models, the dispersion parameter was significantly greater than zero, suggesting use of the negative binomial over a Poisson regression model. All models controlled for the excerpt length and included fixed effects for scoring variability of individual attendings. Standard errors were clustered at the level of the resident. $P < .05$ was considered statistically significant.

Quantitative data analysis was performed using STATA version 15 (StataCorp LLP, College Station, TX).

Results

We reviewed 435 assessments collected between January 2016 and January 2017 (TABLE 2). There were 72 residents and 19 attending anesthesiologists represented, with no differences in gender breakdown between junior (27 of 49 female) and senior (12 of 23 female) residents ($P = .82$). There were 10 male and 9 female anesthesiology attending physicians represented in the data set. From the 435 assessments, 1580 excerpts were generated, with an average of 3.6 excerpts per resident per assessment. Two attendings' assessments were excluded from the regression model because they each submitted only 1 assessment.

Primary Outcome

In the multivariate regression model, female residents were significantly less likely than male residents to be described as agentic after controlling for excerpt length, year of training, and evaluator variability ($\beta = -0.347$; 95% confidence interval [CI] –0.666, –0.028; $P = .033$; FIGURE 1A). Senior residents were more likely to be described as agentic overall ($\beta = 0.702$; 95% CI 0.402–1.002; $P < .001$) compared with junior residents (FIGURE 1B). There was no difference in the distribution of communal or technical feedback by gender or resident level (all $P > .05$).

Agentic Personality Traits Versus Agentic Behavior

Female residents were significantly less likely to be described with agentic *personality traits* than male residents ($\beta = -0.527$; 95% CI –0.923, –0.131; $P = .009$). No significant difference was found between senior and junior residents in this subcategory ($P = .13$). Senior residents were more likely to be described with agentic *behavior traits* compared with

TABLE 1
Coding Schema for Anesthesiology Resident Evaluations[a]

Category	Type and Comments
Agentic (n = 211)	Praise versus criticism
Behavior (n = 112)	"Communicates well with surgical team" "Effectively communicates plan" "Runs the operating room" "Leader in the operating room" "Delegates tasks appropriately" "Ready to take on a challenge" "Decisive" "Self-motivated/self-directed learner" "Stays on top of things" "Always ready for the next step" "Efficient" "Competent"
Personality (n = 99)	"Calm under pressure" "Poised" "Goal-oriented" "Determined" "Resourceful" "Self-sufficient" "Proactive" "Independent" "Confident" "Intelligent" "Bright"
Communal (n = 603)	
Behavior (n = 304)	"Concern for people" "Taught medical students" "Gets along well with others" "Team player" "Good bedside manner" "Works hard/good work ethic" "Asks good questions" "Takes criticism well" "Learns from feedback" "Open to learning"
Personality (n = 299)	"Warm" "Comforting" "Pleasant" "Kind" "Caring/cared for" "Compassionate" "Easy to get along with" "Easy to work with" "Good attitude" "Great to work with/pleasure to work with" "Eager/enthusiastic" "Great energy level/energetic" "Thoughtful/conscientious" "Meticulous/detail oriented"
Technical (n = 597)	
Behavior (n = 597)	"Adapts well to changing circumstances" "Dynamic thinker" "Making appropriate progress" "Upper tier of peer group" "Functions well independently" "Good anesthetic plan" "Situational awareness" "Appreciated complex aspects of case" "Organization (sequence of events, well-prepared for cases, prepared, organized)" "Task (procedures: intravenous therapy, art lines, central lines, intubation, spinals, task completion)" "Knowledge fund/reads on relevant issues"
Other (n = 169)	Praise versus criticism (as appropriate)
Miscellaneous (n = 139)	"Great job" "Top notch" "Did well" "Solid" "Fantastic"
Personality (n = 10)	"Wise beyond years" "Excellent perspective" "Serious"
Confidence/presence (n = 20)	"Working on confidence" "Lacks confidence" "Can be tentative in decision-making" "Passive/meek"

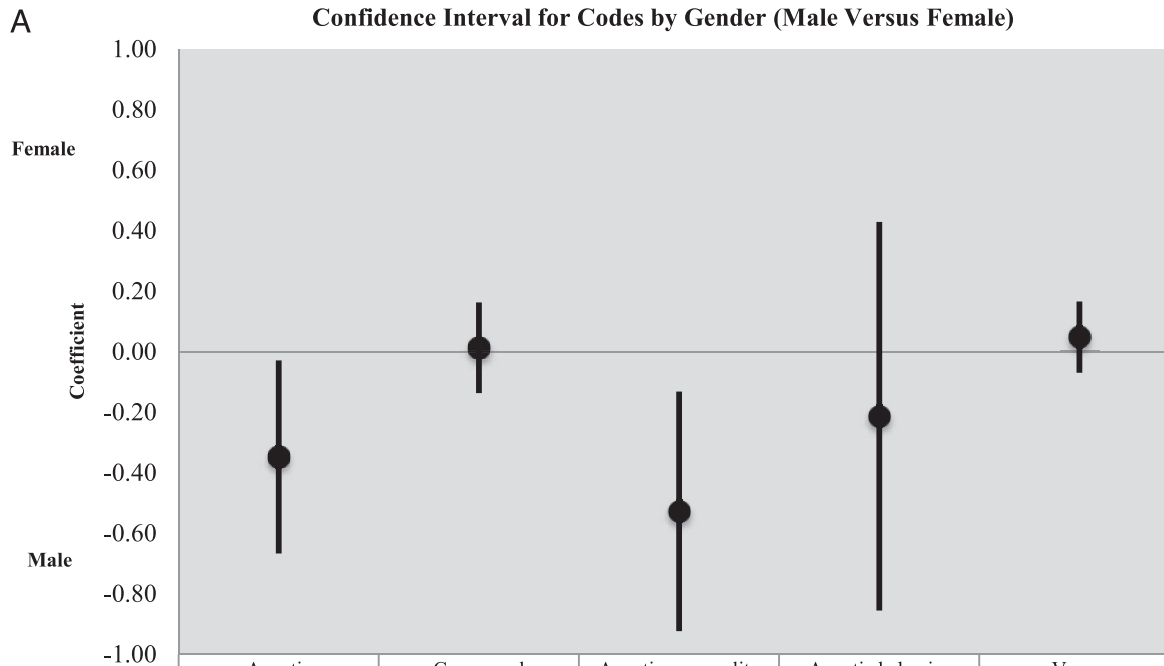
[a] Each thematic excerpt received 1 of each of the following codes: category (agentic, technical, communal, or other); subcategory (agentic behavior, agentic personality, communal behavior, communal personality, technical behavior, other miscellaneous, other personality, other confidence/presence); praise or criticism as appropriate; and vague or specific. There were 1580 excerpts. Additionally, each excerpt was coded as vague or specific (N = 1580).

junior residents ($\beta = 1.055$; 95% CI 0.523–1.588; $P < .001$). No difference was found between male and female residents in this subcategory ($P = .52$).

Interaction Between Gender and Year of Training

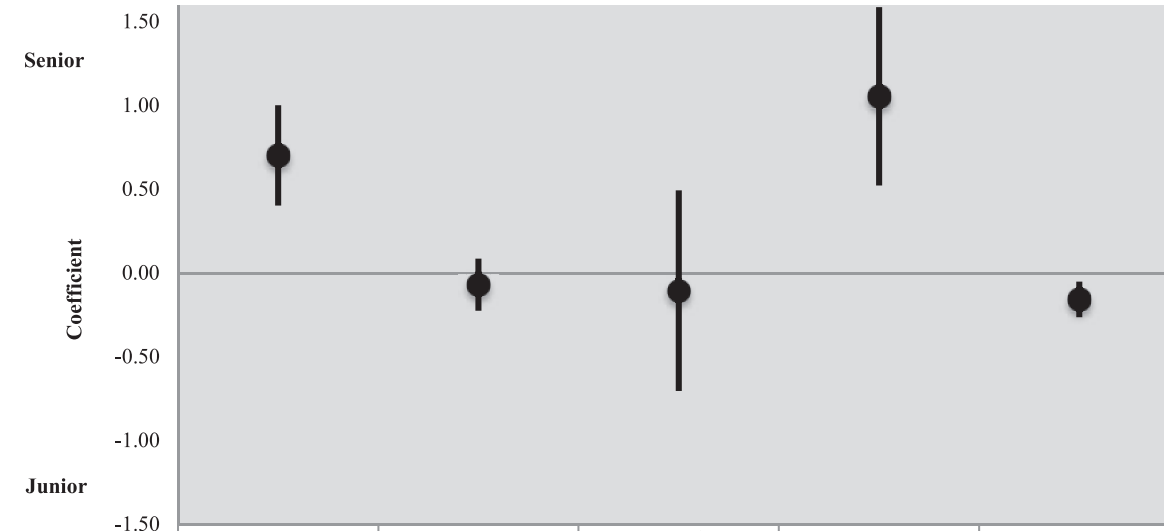
We tested an interaction variable between resident gender and PGY level to examine whether gender played a role in evaluations as training progressed. The interaction coefficient was statistically significant ($\beta = 0.704$; 95% CI 0.084–1.324; $P = .026$), indicating an increased utilization of agentic codes for senior female residents.

A



	Agentic (P = .033)	Communal (P = .86)	Agentic personality (P = .009)	Agentic behavior (P = .52)	Vague (P = .42)
Conf lower bound	-0.67	-0.14	-0.92	-0.86	-0.07
Conf upper bound	-0.03	0.16	-0.13	0.43	0.17
Interval	-0.35	0.01	-0.53	-0.21	0.05

B



	Agentic (P < .001)	Communal (P = .38)	Agentic personality (P = .73)	Agentic behavior (P < .001)	Vague (P = .004)
Conf lower bound	0.40	-0.23	-0.70	0.52	-0.26
Conf upper bound	1.00	0.09	0.49	1.59	-0.05
Interval	0.70	-0.07	-0.11	1.06	-0.16

FIGURE 1A

Confidence Intervals for All Categories by Gender

Note: Confidence intervals are for codes by gender from negative binomial logistic regression. Coefficients are graphed by category (resident) and subcategory (male versus female). All models were controlled for excerpt length, postgraduate year, and resident gender, and included fixed effects for scoring variability of each individual attending. Standard errors were clustered at the level of the resident.

TABLE 2
Demographic Information

Baseline Characteristics	Female		Male		Total	
	n	No. of Evaluations	n	No. of Evaluations	n	No. of Evaluations
Resident anesthesiologist (n = 72)						
Junior cohort	27	126	22	128	49	254
Senior cohort	12	123	11	58	23	181
Total	39	249	33	186	72	435
Attending anesthesiologist (n = 19)	9	205	10	230	19	435

Vague Versus Specific Feedback

Senior residents were less likely to receive vague feedback ($\beta = -0.158$; 95% CI –0.264, –0.051; $P = .004$; FIGURE 1B) than junior residents, with no difference between genders. There was no difference for specific feedback.

Predicting Number of Agentic Comments

FIGURE 2 illustrates the predicted number of agentic comments for a given resident out of a set of 10 assessments, factoring in an interaction between training year and gender. Among junior residents, men would receive 1.8 agentic comments compared with 0.9 agentic comments for women, and among senior residents, men would receive 2.7 agentic comments compared with 2.6 agentic comments for women. There were no differences between groups for communal and technical comment point estimates.

Discussion

Overall, our findings show that female anesthesiology residents were described as agentic less often than male residents. However, when examining subgroups of residents, senior female residents received more agentic feedback compared with junior female residents, and there was no difference in agentic descriptors between senior male and female residents. Identifying these differences in leadership language in resident assessments may help to inform the creation and timing of interventions for faculty assessment training as well as aid the formative development of female residents.

Differences in agentic feedback were driven predominantly by greater use of agentic personality traits in the descriptions of male residents. Words such as "confident," "competent," and "efficient" may be considered more inherent, or fixed, to the individual. These terms are consistent with the concept of a stereotypical "leadership prototype" and are commonly associated with masculine stereotypes in Western culture.[4,19,20] It may not be surprising, then, that junior residents, who have not yet distinguished themselves clinically, are more susceptible to stereotype bias by evaluators. This can be extended to increased vague feedback given to junior residents, which is more likely to be based in generalizations as a result of junior residents being less familiar to evaluators and having fewer observed examples of clinical competence.

The increase in agentic codes observed in the senior cohort was driven predominantly by greater use of agentic behavior descriptions among this group for male and female residents. This category contains behaviors that are more likely to be learned throughout residency training, rather than personality traits that may be intrinsic, vague, and prone to bias. We speculate that as anesthesiology residents develop clinically and interact more frequently with faculty, evaluations will move away from stereotype bias and focus more on demonstrated skills.

The findings in this study support existing literature in academic medicine identifying biases based on stereotype, particularly when evaluating women

←

FIGURE 1B
Confidence Intervals for All Categories by Year

Note: Confidence intervals are for codes by resident year from negative binomial logistic regression. Coefficients are graphed by category (year) and subcategory (senior versus junior). All models were controlled for excerpt length, postgraduate year, and resident gender, and included fixed effects for scoring variability of each individual attending. Standard errors were clustered at the level of the resident.

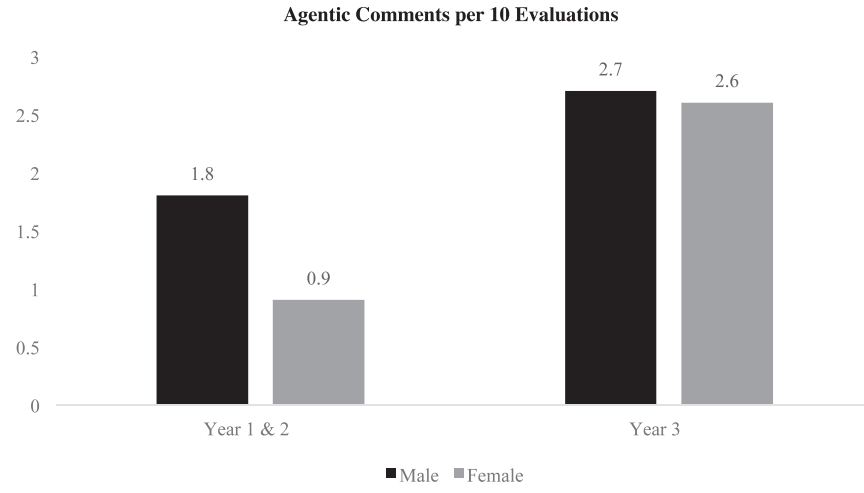


FIGURE 2
Point Estimates by Gender and Year

Note: Per 10 evaluations, estimated number of agentic comments by year and gender. The interaction between resident gender and year suggests that female residents were increasingly likely to receive agentic comments as senior residents compared with more junior female residents ($P = .026$).

for institutional leadership positions.[21,22] Agentic feedback is important given its association with the assessment of leadership capability.[6] Interventions that address this gap in perception of agentic qualities in female residents, particularly related to personality versus behavior-driven feedback,[23] may be best when done early in residency. For individual evaluators of junior residents, a residency program may create a standardized rubric or set of ideal qualities from which they base their written feedback.

This study helps shed light on possible gender stereotype bias applied to residents in their assessments, particularly in the junior stages of residency. However, the agentic-communal framework used for assessing leadership traits is one that is applied most often to traditional nonclinical leadership roles, such as a chief executive officer, vice president, or department chair. There is an emerging body of literature that describes the nature of team leadership in health care action teams that function under complex, dynamic, and time-pressured conditions, such as a code team or intraoperative team. In this context, the nature of leadership is more fluid and task-oriented.[24,25] More research is needed to determine whether this framework provides a more clinically relevant set of criteria for assessing resident progress throughout residency.

This study has several limitations. First, it was conducted at a single center, which decreases generalizability. Second, we were restricted to submitted assessments and therefore lacked complete context (ie, details of the attending-resident interaction). Third, certain characteristics were not available or were removed to preserve anonymity (eg, age, race, and ethnicity of residents and attendings), which may affect feedback.[26,27] Fourth, this study was not longitudinal; therefore, conclusions cannot be drawn regarding how resident evaluations may change over time, and the particular findings may be cohort-specific rather than related to training year. Finally, selection bias cannot be excluded as senior residents chose to rotate at the site while junior residents are generally assigned rotations. We do not know if the sample of senior residents possesses different traits compared with their peers.

Based on the findings of this initial project, further research is needed to better understand how gender bias in resident evaluations may change during the course of training. Additionally, it would be helpful to explore opportunities to intervene as well as the timing of these interventions to affect how residents are perceived and described in their assessments.

Conclusion

In this cross-sectional study, female anesthesiology residents are less likely to be described using "leadership language" in faculty assessments during the first 2 years of training. This gender gap is not found among senior residents. Additionally, male junior residents are more often described using leadership personality traits while senior residents, both male and female, are more likely to receive feedback describing leadership behaviors.

References

1. Lautenberger DM, Dandar VM, Raezer CL, Sloane RA. *The State of Women in Academic Medicine: The Pipeline and Pathway to Leadership*. Washington, DC: Association of American Medical Colleges; 2014. https://members.aamc.org/eweb/upload/The State of Women in Academic Medicine 2013-2014 FINAL.pdf. Accessed December 3, 2018.
2. Association of American Medical Colleges. *The State of Women in Academic Medicine: The Pipeline and Pathways to Leadership, 2015–2016*. https://www.aamc.org/members/gwims/statistics. Accessed December 3, 2018.
3. Edmunds LD, Ovseiko PV, Shepperd S, Greenhalgh T, Frith P, Roberts NW, et al. Why do women choose or reject careers in academic medicine? A narrative review of empirical evidence. *Lancet*. 2016;388(10062):2948–2958. doi:10.1016/S0140-6736(15)01091-0.
4. Koenig AM, Eagly AH, Mitchell AA, Ristikari T. Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychol Bull*. 2011;137(4):616–642. doi:10.1037/a0023557.
5. Salles A, Milam L, Cohen G, Mueller C. The relationship between perceived gender judgment and well-being among surgical residents. *Am J Surg*. 2018;215(2):233–237. doi:10.1016/j.amjsurg.2017.08.049.
6. 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.
7. Dutt K, Pfaff DL, Bernstein AF, Dillard JS, Block CJ. Gender differences in recommendation letters for postdoctoral fellowships in geoscience. *Nat Geosci*. 2016;9(11):805–808. doi:10.1038/ngeo2819.
8. Abele AE, Wojciszke B. Communal and agentic content in social cognition: a dual perspective model. *Adv Exp Soc Psychol*. 2014;50:195–255. doi:10.1016/B978-0-12-800284-1.00004-7.
9. Abele AE. The dynamics of masculine-agentic and feminine-communal traits: findings from a prospective study. *J Pers Soc Psychol*. 2003;85(4):768–776. doi:10.1037/0022-3514.85.4.768.
10. Egal AH, Karau SJ. Role congruity of prejudice toward female leaders. *Psychol Rev*. 2002;109(3):573–598. doi:10.1037/0033-295X.109.3.573.
11. Leonard R. Theorizing the relationship between agency and communion. *Theory Psychol*. 1997;7(6):823–835. doi:10.1177/0959354397076005.
12. Bakan D. *The Duality of Human Existence: Isolation and Communion in Western Man*. 1st ed. Boston, MA: Beacon Press; 1966.
13. Weiss M, Kolbe M, Grote G, Dambach M, Marty A, Spahn DR, et al. Agency and communion predict speaking up in acute care teams. *Small Group Res*. 2014;45(3):290–313. doi:10.1177/1046496414531495.
14. Hoyt CL, Burnette JL. Gender bias in leader evaluations: merging implicit theories and role congruity perspectives. *Personal Soc Psychol Bull*. 2013;39(10):1306–1319. doi:10.1177/0146167213493643.
15. Chiu C, Hong Y, Dweck CS. Lay dispositionism and implicit theories of personality. *J Pers Soc Psychol*. 1997;73(1):19–30. doi:10.1037/0022-3514.73.1.19.
16. Hruschka DJ, Schwartz D, St. John DC, Picone-Decaro E, Jenkins RA, Carey JW. Reliability in coding open-ended data: lessons learned from HIV behavioral research. *Field Methods*. 2004;16(3):307–331. doi:10.1177/1525822X04266540.
17. De Vries H, Elliott MN, Kanouse DE, Teleki SS. Using pooled kappa to summarize interrater agreement across many items. *Field Methods*. 2008;20(3):272–282. doi:10.1177/1525822X08317166.
18. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37–46. doi:10.1177/001316446002000104.
19. Forsyth DR, Nye JL. Seeing and being a leader: the perceptual, cognitive, and interpersonal roots of conferred influence. *Leadersh Crossroads Leadersh Psychol*. 2008;1:116–131.
20. Eagly AH, Carli LL. *Through the Labyrinth: The Truth About How Women Become Leaders*. Boston, MA: Harvard Business Review Press; 2007.
21. Dayal A, O'Connor DM, Qadri U, Arora VM. Comparison of male vs female resident milestone evaluations by faculty during emergency medicine residency training. *JAMA Intern Med*. 2017;177(5):651–657. doi:10.1001/jamainternmed.2016.9616.
22. Mueller AS, Jenkins TM, Osborne M, Dayal A, O'Connor DM, Arora VM. Gender differences in attending physicians' feedback to residents: a qualitative analysis. *J Grad Med Educ*. 2017;9(5):577–585. doi:10.4300/JGME-D-17-00126.1.
23. Cecchi-Dimeglio P. How gender bias corrupts performance reviews, and what to do about it. *Harvard Business Review*. April 12, 2017. https://hbr.org/2017/04/how-gender-bias-corrupts-performance-reviews-and-what-to-do-about-it. Accessed December 3, 2018.
24. Rosenman ED, Ilgen JS, Shandro JR, Harper AL, Fernandez R. A systematic review of tools used to assess team leadership in health care action teams. *Acad Med*. 2015;90(10):1408–1422. doi:10.1097/ACM.0000000000000848.
25. Morgeson FP, DeRue DS, Karam EP. Leadership in teams: a functional approach to understanding

leadership structures and processes. *J Manage*. 2010;36(1):5–39. doi:10.1177/0149206309347376.
26. Rosette AS, Leonardelli GJ, Phillips KW. The white standard: racial bias in leader categorization. *J Appl Psychol*. 2008;93(4):758–777. doi:10.1037/0021-9010.93.4.758.
27. Rosette AS, Livingston RW. Failure is not an option for Black women: effects of organizational performance on leaders with single versus dual-subordinate identities. *J Exp Soc Psychol*. 2012;48(5):1162–1167. doi:10.1016/j.jesp.2012.05.002.

All authors are with Stanford University School of Medicine. **Nicole Arkin, MD,** is a Resident Physician (PGY-4), Department of Anesthesiology, Perioperative, and Pain Medicine; **Cara Lai, BS,** is a Medical Student (MS-3); **Larissa Miyachi Kiwakyou, MD,** is a Resident Physician (PGY-1), Department of Anesthesiology, Perioperative, and Pain Medicine; **Gregory Milo Lochbaum, MD,** is Clinical Assistant Professor, Department of Anesthesiology, Perioperative, and Pain Medicine, and Staff Physician, Anesthesiology and Perioperative Care Service, Veterans Affairs Palo Alto Health Care System; **Audrey Shafer, MD,** is Professor, Department of Anesthesiology, Perioperative, and Pain Medicine, and Staff Physician, Anesthesiology and Perioperative Care Service, Veterans Affairs Palo Alto Health Care System; **Steven K. Howard, MD,** is Professor, Department of Anesthesiology, Perioperative, and Pain Medicine, and Staff Physician, Anesthesiology and Perioperative Care Service, Veterans Affairs Palo Alto Health Care System; **Edward R. Mariano, MD, MAS,** is Professor, Department of Anesthesiology, Perioperative, and Pain Medicine, and Chief, Anesthesiology and Perioperative Care Service, Veterans Affairs Palo Alto Health Care System; and **Magali Fassiotto, PhD,** is Assistant Dean for Faculty Development and Diversity.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

The authors would like to thank Vivianne Tawfik, MD, PhD, and the research elective team in the Department of Anesthesiology, Perioperative, and Pain Medicine at Stanford University School of Medicine for providing the necessary research time to explore this important topic. The authors would also like to thank Ronald Pearl, MD, PhD, Department Chair, and the anesthesiology faculty for training and supporting residents as they attempt to answer difficult questions and pursue their research interests.

Corresponding author: Nicole Arkin, MD, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, 443.326.8544, narkin@stanford.edu

Received May 14, 2018; revisions received September 16, 2018, and October 26, 2018; accepted October 26, 2018.