

Gender Differences in Patient Perceptions of Physicians' Communal Traits and the Impact on Physician Evaluations

Heidi Chen, BS,¹ Emma Pierson, PhD,¹ Sonja Schmer-Galunder, MS,² Jonathan Altamirano, MS,³ Dan Jurafsky, PhD,^{1,4} Jure Leskovec, PhD,¹ Magali Fassiotto, PhD,³ and Nishita Kothary, MD⁵

Abstract

Background: Communal traits, such as empathy, warmth, and consensus-building, are not highly valued in the medical hierarchy. Devaluing communal traits is potentially harmful for two reasons. First, data suggest that patients may prefer when physicians show communal traits. Second, if female physicians are more likely to be perceived as communal, devaluing communal traits may increase the gender inequity already prevalent in medicine. We test for both these effects.

Materials and Methods: This study analyzed 22,431 Press Ganey outpatient surveys assessing 480 physicians collected from 2016 to 2017 at a large tertiary hospital. The surveys asked patients to provide qualitative comments and quantitative Likert-scale ratings assessing physician effectiveness. We coded whether patients described physicians with “communal” language using a validated word scale derived from previous work. We used multivariate logistic regressions to assess whether (1) patients were more likely to describe female physicians using communal language and (2) patients gave higher quantitative ratings to physicians they described with communal language, when controlling for physician, patient, and comment characteristics.

Results: Female physicians had higher odds of being described with communal language than male physicians (odds ratio 1.29, 95% confidence interval 1.18–1.40, $p < 0.001$). In addition, patients gave higher quantitative ratings to physicians they described with communal language. These results were robust to inclusion of controls.

Conclusions: Female physicians are more likely to be perceived as communal. Being perceived as communal is associated with higher quantitative ratings, including likelihood to recommend. Our study indicates a need to reevaluate what types of behaviors academic hospitals reward in their physicians.

Keywords: gender differences, female physicians, language, quality care, communal traits, Press Ganey

Introduction

IN 1995, KATHLEEN HALL JAMIESON proposed five double binds that professional women face when they violate gender perceptions and societal prescriptions.¹ The Same-ness/Difference double bind describes societal biases that punish women for agentic traits (*e.g.*, strong, logical, and independent) that are rewarded in men while simultaneously devaluing communal traits (empathy, nurturing, consensus-building) which women often foster.¹ Academic medicine is no exception, and a large body of research has explored this dearth of women in leadership in the field.^{2–4} Institutional

emphasis on agentic traits in academic medicine is not only front and center but also resides squarely in the “think manager (leader), think male” mindset.^{5–8}

Yet, communal characteristics are the cornerstones of medicine.⁹ Patients value physicians who are empathetic and personable and value shared decision-making, all canonical communal traits.^{9–12} Patient-centered medicine and empathetic communication improve patient compliance to prescribed medical treatment and influence recovery and clinical outcome.^{13,14} In other words, the very traits that are negatively associated with advancement in academic medicine^{11,12} are also those valued by patients. These patient

¹Department of Computer Science, Stanford University, Stanford, California, USA.

²Smart Information Flow Technologies, Minneapolis, Minnesota, USA.

³Office of Faculty Development and Diversity, Stanford School of Medicine, Stanford, California, USA.

⁴Department of Linguistics, Stanford University, Stanford, California, USA.

⁵Department of Radiology, Stanford School of Medicine, Stanford, California, USA.

evaluations are playing an increasing role in MD performance evaluation and subsequent advancement, and the role of gender in this interplay between physician communal traits and patient expectations is increasingly relevant in the evolving doctor-patient relationship. Whether patients too view physicians through the lens of gender stereotype and thus ascribe communal traits more often to female physicians remains to be investigated. Furthermore, the correlation (positive or negative) between the presence of these communal traits and patient satisfaction scores remains an important and unanswered question.

In this study, we analyze (1) whether patient-ascribed communal traits had a gendered prevalence (*i.e.*, whether words such as “compassionate” or “warm” were used more frequently for male vs. female physicians) and (2) whether physicians described using communal language received higher or lower quantitative scores. Demonstrating the value patients place on communal traits, as reflected by their quantitative physician assessments, could serve as a call to reform a field that is notorious for devaluing communal behavior.

Materials and Methods

The study was granted an exemption by the Institutional Research Board Review for the Protection of Human Subjects. Patient experience and satisfaction with their physicians is commonly captured on postvisit surveys. The Press-Ganey patient satisfaction survey is one such scientifically rigorous and validated tool that measures patient satisfaction using qualitative free-text and a quantitative 5-point score, which is used for performance improvement and as an individual yard-stick for physicians.¹⁵ Since language is one of the most powerful means through which society continues to perpetrate conscious and subconscious biases, we investigate whether (1) patients are more likely to describe female physicians using communal language, and (2) physicians described with communal language receive higher or lower quantitative Press-Ganey scores.

Data

The study was granted an exemption by the Institutional Research Board Review for the Protection of Human Subjects. Deidentified Press Ganey surveys of outpatient visits at the hospital were collected over a period of 2 years (2016–2017). As per hospital policy, all patients registered in the hospital’s online patient portal receive a survey following an outpatient visit. In addition, each month five randomly selected patients per physician are mailed a paper version of the survey. The total response rate is 25%.

Each survey included both a qualitative comments section and a quantitative ratings section. Each of these sections included multiple subsections where patients were asked to assess specific aspects of the visit. The qualitative comments section asked patients to assess *Access*, *Moving Through Your Visit*, *Nurse/Assistant*, *Personal Issues*, *Care Provider*, and provide an *Overall Assessment*. We removed comments about *Access*, *Moving Through Your Visit*, and *Nurse/Assistant* because they were not directly relevant to the patient-physician interaction. We removed comments not made in English (~3% of the original distribution) because our measures of communal language relied on English word lists.

The quantitative ratings section asked patients to assess physicians on a number of dimensions on a 1 to 5 Likert scale, with five denoting more positive experiences. We assessed the following dimensions: (1) *Care provider’s efforts to include you in decisions about your treatment*, (2) *Concern the care provider showed for your questions or worries*, (3) *Your confidence in this care provider*, (4) *Friendliness/courtesy of the care provider*, (5) *Likelihood of your recommending this care provider to others*, and (6) *Likelihood of your recommending our practice to others*. We place particular emphasis on the “likelihood of your recommending this care provider to others,” as it is used by most health care organizations as a primary metric in determining physician performance.¹⁶

We filtered for surveys, in which both the quantitative and qualitative portions of the data were available, leaving us with 22,431 surveys of 480 physicians. Each observation in the dataset thus consisted of one survey for one patient visit; a single physician could be assessed in many surveys. Physicians were identified with an anonymous identifier. Patient data were anonymous and included no patient identifier. Our dataset also included demographic information about both the physician and patient: physician gender, age, ethnicity, race, rank, and medical specialty, and patient gender, age, and race. Table 1 provides summary statistics for physicians and patients.

Language processing

We defined a survey response as containing communal language using a word scale based on a previous study of communal language.¹⁷ As our domain differed somewhat from that in the original study—letters of recommendation—we modified the scale to remove words which were not salient to our domain or falsely identified comments. The final scale we used identified text as communal if it contained any of the words or prefixes listed in Figure 1.

We confirmed, by hand inspection of a random sample of surveys in our dataset, that surveys identified as containing communal language were consistent with the intended scale. For our analysis, we used a binary communal variable defined to be 1 if a comment contained any communal words and 0 otherwise. (Results were similar when analysis was performed instead using the count of communal words in the comment.)

Regression analysis

We analyzed (1) whether patients were more likely to describe female physicians using communal language and (2) whether physicians described using communal language received higher or lower quantitative ratings.

In all regressions, we included *physician controls* (age, rank, medical specialty, and race/ethnicity); *patient controls* (gender, age, and race); and *comment controls* (both comment length as well as which qualitative survey questions were being answered). To determine whether patients were more likely to describe female physicians using communal language, we used physician gender as the independent variable and presence of communal language as the dependent variable. To determine whether communal language was associated with higher quantitative ratings, we used presence

TABLE 1. SUMMARY OF PHYSICIAN AND PATIENT STATISTICS FOR THE DATASET

<i>Physician demographics</i>	
No. of physicians	480
% male	56%
Mean age (SD)	46 (11)
Race/Ethnicity	
% White	51%
% Asian	33%
% Hispanic/Latino	4%
% Black	1%
% Native Hawaiian/Pacific Islander	<1%
% Other/mixed/declined to state	10%
Academic rank	
% Clinical instructor	7%
% Clinical assistant professor	36%
% Clinical associate professor	11%
% Clinical professor	6%
% Assistant professor	11%
% Associate professor	10%
% Professor	18%
Medical fields included	Medicine, orthopedic surgery, neurology, ophthalmology, dermatology, surgery, otolaryngology, neurosurgery, obstetrics and gynecology, psychiatry, urology, cardiothoracic surgery, pediatrics, cardiovascular medicine
Mean surveys per physician	47
Median surveys per physician	36
<i>Patient demographics</i>	
% Male	40%
Mean age (SD)	62 (16)
Race ^a	
% White	65%
% Asian	12%
% Black	2%
% Unknown/Other/Patient Refused	21%

^aEthnicity demographic information was not available on the survey.
SD, standard deviation.

of communal language as the independent variable and quantitative rating as the dependent variable.

Because patients gave very high quantitative ratings to physicians (all quantitative dimensions we examined were at least 4.6/5 on average, with 5s given in at least 85% of cases), we evaluated ratings dichotomously with the commonly used Top Box metric defined by Press Ganey. A physician is given a 1 in Top Box if they receive the highest possible score on that question (in this case, a 5/5), and a 0 otherwise. This scoring is used to calculate all Consumer Assessment of

Healthcare Providers and Systems metrics and is what institutions use to evaluate and reward quality treatment.¹⁵ Although small percentage differences in Top Box may seem insignificant, these variations lead to substantial differences in physician rankings. According to recently published Press Ganey statistics, physicians with 80% Top Box ratings in “Likelihood of your recommending this care provider to others” ranked in the 24th percentile of physicians, whereas those with 90% Top Box ratings ranked in the 94th percentile.¹⁸

To account for multiple correlated observations for each physician, we computed cluster-robust standard errors, clustering at the physician level. Because all dependent variables were binary, we used logistic regressions for all analyses; regression coefficients and odds ratios (ORs) are reported below.

Results

Female physicians were more likely to be described using communal language. Overall, 17% of comments contained communal language. Without inclusion of controls, the odds of female physicians being described with communal language were higher than for their male counterparts (OR = 1.29, 95%

Words: genial, likeable, union, solidarity, affectionate, fond, kind, sympathetic, sensitive, nurturing, communal, responsive, considerate, agreeable, interpersonal, warm, caring, tactful, sweet, friendly, accepting, supportive, nice, pleasant, compassionate, encouraging, welcoming, gentle

Prefixes: tolera-, cooperat-, sympath-, nurtur-, interpersona-, intima-, connec-, agreeable-

FIG. 1. Communal language word scale.

TABLE 2. COMMUNAL LANGUAGE IS ASSOCIATED WITH HIGHER QUANTITATIVE SCORES

<i>Quantitative rating</i>	<i>Noncommunal comments (%)</i>	<i>Communal comments (%)</i>	<i>Adjusted OR (95% CI)</i>
Care provider's efforts to include you in decisions about treatment (5/5)	82	89	2.52 (2.21–2.87)
Concern care provider showed for your questions or worries (5/5)	82	89	2.39 (2.10–2.73)
Your confidence in care provider (5/5)	83	90	2.64 (2.31–3.00)
Care provider friendliness/courtesy (5/5)	86	92	2.46 (2.14–2.82)
Likelihood of your recommending care provider to others (5/5)	83	90	2.77 (2.43–3.15)
Likelihood of your recommending our practice to others (5/5)	81	88	2.59 (2.21–2.87)

The independent variable is the presence of communal language, and the dependent variable is the Top Box score, a dichotomous variable indicating whether the physician received a perfect rating of 5/5 for each quantitative dimension. The first two numerical columns report the percentage of comments which received perfect 5/5 scores with/without communal language. The final column reports the OR after adjusting for patient, physician, and comment controls, with 95% confidence interval in parentheses; all results are statistically significant ($p < 0.001$). Results remain statistically significant and substantively large (all ORs > 1.7 , $p < 0.001$) without controls.

CI, confidence interval; OR, odds ratio.

confidence interval [95% CI] = 1.18–1.40, $p < 0.001$). With inclusion of physician, patient, and comment controls, the odds of female physicians being described with communal language remained significantly higher (OR = 1.15, 95% CI = 1.05–1.26, $p = 0.002$).

Physicians described with communal language were significantly more likely to receive perfect quantitative ratings of 5/5 (Tables 2 and 3) across all the quantitative dimensions when controlling for physician, patient, and comment characteristics, with ORs of at least 2.39 and $p < 0.001$ in all cases. Of particular importance to patient care were the fields for “Care provider’s efforts to include you in decisions about your treatment” (OR = 2.52, 95% CI = 2.21–2.87), “Concern the care provider showed for your questions or worries” (OR = 2.39, 95% CI = 2.10–2.73), “Your confidence in this care provider” (OR = 2.64, 95% CI = 2.31–3.00), and “Likelihood of your recommending this care provider to others” (OR = 2.77, 95% CI = 2.43–3.15). In this last field, “Likelihood of your recommending this care provider to others,” 81% of physicians who were described without communal language received 5/5, while 88% of physicians described with communal language received 5/5.

This discrepancy represents a notable increase from the 28th percentile to the 84th percentile in physician Press Ganey rankings for that survey question.¹⁸

We also assessed whether the positive association between communal language and quantitative ratings was larger or smaller for women: that is, whether the gender-communal language interaction term was statistically significant. We found no statistically significant interaction; there was no evidence that physicians of either gender experienced greater returns to communal language. We also found no significant difference in effect sizes between surgical and nonsurgical fields.

We additionally assessed whether female physicians earned significantly higher or lower quantitative ratings than male physicians across the board, using a regression where physician gender was the independent variable and quantitative rating was the dependent variable. We found no statistically significant differences.

Discussion

This study highlights three important findings. First, the results of our study indicate that physicians with communal traits

TABLE 3. ADDITIONAL BREAKDOWN OF QUANTITATIVE SCORES BY GENDER AND PRESENCE OF COMMUNAL LANGUAGE

<i>Quantitative rating</i>	<i>Male, noncommunal (%)</i>	<i>Male, communal (%)</i>	<i>Female, noncommunal (%)</i>	<i>Female, communal (%)</i>
Care provider's efforts to include you in decisions about treatment (5/5)	82	89	82	90
Concern care provider showed for your questions or worries (5/5)	82	88	83	90
Your confidence in care provider (5/5)	84	90	83	90
Care provider friendliness/courtesy (5/5)	86	91	86	92
Likelihood of your recommending care provider to others (5/5)	83	90	82	91
Likelihood of your recommending our practice to others (5/5)	82	88	79	88

As in Table 2, the independent variable is the presence of communal language, and the dependent variable is the Top Box score. Scores are broken down into categories based on physician gender and the presence of communal language in their patients' comments. Both male and female physicians described with communal language experienced higher quantitative results than when described without communal language. We found no statistically significant difference in regressions run on the gender-language interaction term: there was no evidence that physicians of either gender experienced greater returns from communal language in their quantitative ratings.

receive higher quantitative ratings. Second, among these ratings, patients' appreciation for communal traits were associated with the important, commonly used "likelihood to recommend" yardstick.¹⁶ Finally, patients perceive a gender difference in communal traits and are more likely to identify these traits in female physicians, even after controlling for physician and patient demographics. Whether this pattern stems from unconscious gender stereotyping by patients or from actual differences in female physicians' behavior is difficult to tease out from descriptive narratives. Despite these findings, overall scores between men and women physicians were statistically similar. This may be due to two reasons. First, statistically combining two modest correlations (1) physician gender and communal language and (2) communal language and quantitative scores—results in a combined correlation, which is less likely to be statistically significant. Second, several other factors influence quantitative scores besides communal traits, some of which may put women at a disadvantage.^{19,20}

Our study builds upon previous work suggesting that the doctor-patient relationship has changed in recent years. While in the past the relationship could be viewed as that of medical paternalism where the "doctor knew best," shifting societal values, subspecialization, and an increasingly complex health care system have made shared decision-making and patient autonomy more important.^{10,12,21} Our study highlights how this shift measurably manifests in physician evaluations by associating patient perception of communal traits with higher likelihood of recommending a care provider, an important factor influencing physician bonuses and career advancement. We additionally identify a correlation between perception of these communal traits and physician gender. This measurable gap in perception is supported by previous studies, which suggest that female physicians are more likely to be perceived as building collaborative models of patient-physician relationships than their male colleagues.^{9,22–24} Female physicians may spend more time with their patients,^{23,24} investigate social and psychological issues in more depth and facilitate patient participation more effectively.^{9,23,24}

Given the importance of communal skills in clinical practice, it may seem surprising that academic medicine continues to reward agentic traits while overlooking, or worse, devaluing, communal traits that are overwhelmingly valued by patients.^{5,25–27} Cultural stereotypes in medicine, perpetrated through language and communication, characterize women as communal or the labor class, describing them as hard-working, dependable, and group-oriented, while men appear to have innate leadership skills, and are deemed brilliant and independent thinkers.^{6–8}

Thus, female physicians in academic medicine face a difficult dilemma. Recognizing that their perceptible communal skillset could actually hinder their professional advancement, female faculty have described attempts to mask or counter stereotypical assumptions.^{28,29} On the contrary, agentic women go through great lengths to avoid violating the gender stereotype (*e.g.*, being perceived as tough, decisive), lest they be stigmatized by both men and women.³⁰

It could be argued that medicine requires quick thinking and decisiveness. However, while that skillset may be critical in emergent situations, by and large, most nonurgent situations call for empathy, listening, shared decision-making, and consensus building. Therefore, it is high time that medicine consciously recognizes and rewards physicians of any gender

who exhibit and cultivate communal traits that would ultimately deliver healthier, happier patients.

Limitations

This study is a single center, retrospective study, limiting generalization. The data were based on Press Ganey survey results and could represent a potentially biased sample of outpatients who choose to fill out the survey. The data have a skewed distribution, with most surveys rating physicians 5/5. Each survey was treated as an individual patient entry and did not account for repeating patient visits. The results in the study are purely correlational and not causal. Finally, different medical specialties require different approaches beyond an analysis of language in patient care.

Conclusions

Patients identified communal traits more frequently in female physicians than their male counterparts, and physicians perceived as communal had higher odds of receiving 5/5 quantitative ratings, including likelihood to recommend. These results have significant implications for how the medical community evaluates the quality of physicians' medical practices. Further, in-depth linguistic analyses across specialties and across all mediums of patient-physicians communication are needed.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

No funding was received for this article.

References

1. Jamieson KH, Hall K. *Beyond the Double Bind: Women and Leadership*. New York, NY: Oxford University Press on Demand, 1995.
2. Mangurian C, Linos E, Sarkar U, Rodriguez C, Jagsi R. What's holding women in medicine back from leadership. Harvard Business Review Web site. 2018. Available at: <https://hbr.org/2018/06/whats-holding-women-in-medicine-back-from-leadership> Accessed March 2, 2020.
3. Travis E. Academic Medicine Needs More Women Leaders. AAMC. 2018. Available at: <https://www.aamc.org/news-insights/insights/academic-medicine-needs-more-women-leaders> Accessed March 2, 2020.
4. Jagsi R, Means O, Lautenberger D, et al. Women's representation among members and leaders of national medical specialty societies. *Acad Med* 2020;95:1043–1049.
5. Carnes M, Morrissey C, Geller SE. Women's health and women's leadership in academic medicine: Hitting the same glass ceiling? *J Women's Health* 2008;17:1453–1462.
6. Zhuge Y, Kaufman J, Simeone DM, Chen H, Velazquez OC. Is there still a glass ceiling for women in academic surgery? *Ann Surg* 2011;253:637–643.
7. Burgess DJ, Joseph A, Van Ryn M, Carnes M. Does stereotype threat affect women in academic medicine? *Acad Med* 2012;87:506.
8. Kaatz A, Carnes M. Stuck in the out-group: Jennifer can't grow up, Jane's invisible, and Janet's over the hill. *J Womens Health (Larchmt)* 2014;23:481–484.
9. Roter DL, Hall JA. Why physician gender matters in shaping the physician-patient relationship. *J Womens Health* 1998;7:1093–1097.

10. Bensing J, Rimondini M, Visser A. What patients want. *Patient Educ Couns* 2013;90:287–290.
11. Bendapudi NM, Berry LL, Frey KA, Parish JT, Rayburn WL. Patients' perspectives on ideal physician behaviors. *Mayo Clin Proc*. 2006;81:338–344.
12. Deledda G, Moretti F, Rimondini M, Zimmermann C. How patients want their doctor to communicate. A literature review on primary care patients' perspective. *Patient Educ Couns* 2013;90:297–306.
13. Stewart M, Brown JB, Boon H, Galajda J, Meredith L, Sangster M. Evidence on patient-doctor communication. *Cancer* 1999;25:30.
14. Haynes RB, McDonald HP, Garg AX. Helping patients follow prescribed treatment: Clinical applications. *JAMA* 2002;288:2880–2883.
15. Associates PG. Scoring Quick Guide for Quick Reports: Rank, Top Box or Mean Score? 2013. Available at: https://helpandtraining.pressganey.com/lib-docs/default-source/ip-training-resources/Scoring_Quick_Guide_for_Quick_Reports.pdf?sfvrsn=0. Accessed August 19, 2020.
16. Garcia LC, Chung S, Liao L, et al. Comparison of outpatient satisfaction survey scores for asian physicians and non-hispanic white physicians. *JAMA Network Open* 2019; 2:e190027–e190027.
17. Madera JM, Hebl MR, Martin RC. Gender and letters of recommendation for academia: Agentic and communal differences. *J Appl Psychol* 2009;94:1591.
18. Associates PG. Medical practice top box % and ranks, likelihood of recommending cp, national facilities. Press Ganey. 2019. Available at: <https://identity.pressganey.com/> Accessed August 31, 2019.
19. Rogo-Gupta LJ, Haunschild C, Altamirano J, Maldonado YA, Fassiotto M. Physician gender is associated with Press Ganey Patient satisfaction scores in outpatient gynecology. *Womens Health Issues* 2018;28:281–285.
20. Cuddy AJC, Fiske ST, Glick P. When professionals become mothers, warmth doesn't cut the ice. *J Social Issues* 2004; 60:701–718.
21. Waitzkin H. Doctor-patient communication: Clinical implications of social scientific research. *JAMA* 1984;252: 2441–2446.
22. Roter DL, Geller G, Bernhardt BA, Larson SM, Doksum T. Effects of obstetrician gender on communication and patient satisfaction. *Obstet Gynecol* 1999;93:635–641.
23. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Gender in medical encounters: An analysis of physician and patient communication in a primary care setting. *Health Psychol* 1994;13:384.
24. Law S, Britten N. Factors that influence the patient centredness of a consultation. *Br J Gen Pract* 1995;45: 520–524.
25. Tesch BJ, Wood HM, Helwig AL, Nattinger AB. Promotion of women physicians in academic medicine: Glass ceiling or sticky floor? *JAMA* 1995;273:1022–1025.
26. Wright AL, Schwindt LA, Bassford TL, et al. Gender differences in academic advancement: Patterns, causes, and potential solutions in one US College of Medicine. *Acad Med* 2003;78:500–508.
27. Ash AS, Carr PL, Goldstein R, Friedman RH. Compensation and advancement of women in academic medicine: Is there equity? *Am J Ophthalmol* 2004;138:903–904.
28. Drago R, Colbeck CL, Stauffer KD, et al. The avoidance of bias against caregiving the case of academic faculty. *Am Behav Sci* 2006;49:1222–1247.
29. Fassiotto M, Hamel EO, Ku M, et al. Women in academic medicine: Measuring stereotype threat among junior faculty. *J Womens Health* 2016;25:292–298.
30. Eagly AH, Carli LL. *Through the Labyrinth: The Truth About How Women Become Leaders*. Boston, MA: Harvard Business School Press, 2007.

Address correspondence to:

Heidi Chen, BS
 Department of Computer Science
 Stanford University
 353 Serra Mall
 Stanford, CA 94305-6104
 USA

E-mail: hchen7@cs.stanford.edu