

Gender Bias in Collaborative Medical Decision Making: Emergent Evidence

Erik G. Helzer, PhD, Christopher G. Myers, PhD, Christine Fahim, PhD, Kathleen M. Sutcliffe, PhD, and James H. Abernathy, MD, MPH

Abstract

This initial, exploratory study on gender bias in collaborative medical decision making examined the degree to which physicians' reliance on a team member's patient care advice differs as a function of the gender of the advice giver. In 2018, 283 anesthesiologists read a brief, online clinical vignette and were randomly assigned to receive treatment advice from 1 of 8 possible sources (physician or nurse, man or woman, experienced or inexperienced). They then indicated their treatment decision, as well as the degree to which they relied upon the advice given.

The results revealed 2 patterns consistent with gender bias in participants' advice

taking. First, when treatment advice was delivered by an inexperienced physician, participants reported replying significantly more on the advice of a man versus a woman, $F(1,61) = 4.24$, $P = .04$. Second, participants' reliance on the advice of the woman physician was a function of her experience, $F(1,62) = 6.96$, $P = .01$, whereas reliance on the advice of the man physician was not, $F(1,60) = 0.21$, $P = .65$.

These findings suggest women physicians, relative to men, may encounter additional hurdles to performing their jobs, especially at early stages in their careers. These hurdles are rooted in psychological biases of others,

rather than objective features of cases or treatment settings. Cultural stereotypes may shape physicians' information use and decision-making processes (and hinder collaboration), even in contexts that appear to have little to do with social category membership. The authors recommend institutions adopt policies and practices encouraging equal attention to advice, regardless of the source, to help ensure advice taking is a function of information quality rather than the attributes of the advice giver. Such policies and practices may help surface and implement diverse expert perspectives in collaborative medical decision making, promoting better and more effective patient care.

Medical decision making is becoming increasingly collaborative and team based. Although patient care decisions focus on specialized knowledge and evidence, to reach decisions providers must assimilate information from multiple sources, including the opinions and observations of other care team members.¹ In medicine, as in other professional contexts, the inherently social nature of collaborative decision making can invite unwanted influences, such as the activation of stereotypes and cultural biases that affect decision makers' interpretation of and reliance upon information provided

Please see the end of this article for information about the authors.

Correspondence should be addressed to James H. Abernathy, Department of Anesthesiology and Critical Care Medicine, Johns Hopkins Medicine, 1800 Orleans St., Bloomberg Children's Center, Suite 6302, Baltimore, MD 21287; email: jaberna4@jhmi.edu.

Written work prepared by employees of the Federal Government as part of their official duties is, under the U.S. Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the United States Code is not available. As such, copyright does not extend to the contributions of employees of the Federal Government.

Acad Med. 2020;95:1524–1528.

First published online July 14, 2020
doi: 10.1097/ACM.0000000000003590

by other team members on the basis of irrelevant characteristics like gender.

In this article, we share the results of an initial, exploratory study we conducted with anesthesiologists to examine whether reliance on a team member's patient care advice differs as a function of the gender of the advice giver. Based on our results, which indicate the presence of gender bias, we provide insights into organizational practices that can mitigate pernicious decision-making biases in the workplace, with suggestions about how these insights could be applied in the context of collaborative medical decision making.

Background: Cultural Stereotypes and Gender Bias in the Workplace

In decision environments where time is limited, the task is complex, and multiple inputs compete for attention, decision makers often rely upon intuitive, "gut-based" decision making.² Such decision making depends on pattern recognition, that is, quickly fitting new information to existing knowledge structures that have been developed through experience and inherited through culture.^{2,3} For example, through professional experience and trial and error,

medical experts develop cognitive templates that facilitate rapid identification of various medical conditions.² Concurrently, they inherit cultural stereotypes (i.e., beliefs about a person's capabilities or attributes based on their social category membership) that are engrained and reinforced across multiple contexts over the course of a lifetime.⁴ The challenge is that decision makers—even content experts—do not choose which knowledge structures are brought to bear on their intuitive assessments. Although they know the products of their intuitive processing, they generally have limited awareness and control over the processes that shape their use of information and their ultimate decisions.⁵

Cultural stereotypes operate in a largely automatic fashion, distorting people's interpretation and use of new information based upon arbitrary characteristics of the source, such as race or gender.⁴ Stereotypes can be powerful impediments to collaborative work since they are irrelevant to the task at hand and resistant to updating. A host of studies, including psychological experiments using random assignment, have shown the detrimental effects of gender stereotypes on collaborative work. For

example, research shows that women's expertise is discounted relative to men's, and that men's work receives less scrutiny than women's.⁶⁻⁹ In the field of medicine, women continue to be disadvantaged by gender-based disparities in training, feedback, advancement, and pay, and they face more discrimination and harassment at work than men do.^{10,11} Collectively, these findings suggest a potential impediment to collaborative medical decision making: The patient care advice offered by a woman might be discounted relative to the same or lesser quality advice from a man, based only on the advice giver's gender.

To expand the discourse around this important topic and spark new insights about the insidious effects of gender bias in medical decision making, we present the results of our initial study exploring advice taking in a collaborative decision-making task. We decided to focus on a single medical specialty, anesthesiology, which allowed us to develop a decision-making scenario within the domain of expertise of all participants. Anesthesiologists are involved in a wide range of care, so decision making among these expert decision makers has a reach that is both broad and consequential. While our scenario was tailored to our participants' medical expertise, we believe that our findings will be applicable to other fields in medicine, given the generality of the cognitive processes and cultural stereotypes we examined.

Exploring Gender Bias in Advice Taking in a Collaborative Medical Decision-Making Task

Participants and recruitment

Our study participants were 283 anesthesiologists (20% female) with 17.32 mean years of experience. Most (76%) described themselves as attending physicians. The participants were recruited for a brief online study in October and November 2018 from the membership of the Society of Cardiovascular Anesthesiologists via a professional listserv. Each participant was entered into a raffle to win 1 of 5 \$500 gift cards.

The study was reviewed and approved by the Johns Hopkins Homewood Institutional Review Board. Due to the low-risk nature of the study, the

requirement for written informed consent was waived by the institutional review board.

Decision-making scenario and experimental manipulations

In our online experiment, participants read a brief clinical decision-making vignette describing a decision about whether to intubate a patient (Box 1). The vignette, which was intentionally written to produce equivocality in decision making, asked participants to imagine they had arrived at the hospital room of a female patient who had undergone operative coronary revascularization several days earlier and was now experiencing difficulty breathing. After participants learned the relevant vitals, they were randomly assigned to 1 of 8 between-subjects conditions in which they received treatment advice from another care team member within the vignette. The advice remained constant (intubate the patient), but the characteristics of the advice giver differed by gender (man versus woman), experience (2 years versus 15 years), and role (physician versus nurse). For example, in one condition, the advice to intubate came from Angela Smith, MD, a 15-year veteran of the hospital; in another condition, the same advice came from Mark Smith, RN, who is 2 years

post-training. All possible permutations (Angela/Mark, MD/RN, 15-year veteran/2 years post-training) were generated, and each participant received advice from only 1 of the 8 possible advice givers. Aside from the advice giver's name, role, and experience, all other information was held constant between conditions. There was no explicit statement made asking participants to consider any of these pieces of information in making their decision.

After receiving advice, participants were asked to respond to a series of questions, including whether *they* would intubate the patient based on the information presented (yes/no) and their level of confidence in that course of action (1 = not at all confident to 5 = completely confident). Participants were also asked to what degree they relied upon 3 sources of information when thinking about what they would do in this scenario: clinical information about the patient; their own expertise and intuition; and, the focus of our study, the advice of the care team member. These ratings used a 5-point scale (0 = not at all to 4 = very much).

Statistical analysis

To test for bias in participants' decision making, we analyzed data using a general linear model (IBM SPSS Version 25; IBM

Box 1

Clinical Decision-Making Vignette and Between-Subjects Conditions

It's 9:30 PM and you are on call in the main OR. Your pager goes off requesting your assistance with a patient in respiratory distress. You are being summoned to determine whether or not to intubate the patient.

You swallow the last bite of your now-cold dinner, grab the code bag, and run off to the 10th floor. You arrive to the hospital room a little winded and a little disoriented.

You begin to assess the patient's condition: The 64 yo, 110 kg woman in front of you is working to breathe but is not yet exhausted. She is 4 days out from her CABG, spent one night in the ICU, and was transferred to the floor on POD 1. She has an EF of 40%. Her saturations are 92% on a non-rebreather, RR 28, BP 100/65, HR 112 in atrial fibrillation.

As you complete your review, a care team member, [Name, Role], notices you have arrived and rushes in. [Name] is [Experience]. [He/She] asks you what you're thinking about the case, then tells you that [he/she] thinks you should intubate the patient.

Gender	Advice Giver ^a	
	Role	Experience
Angela Smith	MD	a 15-year veteran of the hospital
or	or	or
Mark Smith	RN	2 years post-training

Abbreviations: yo, year old; kg, kilogram; CABG, coronary artery bypass grafting; ICU, intensive care unit; POD, postoperative day 1; EF, ejection fraction; RR, respiratory rate; BP, blood pressure; HR, heart rate.

^aEach participating anesthesiologist received treatment advice from a care team member whose identity was described as 1 of 8 possible combinations of gender (manipulated using names), role, and experience.

Corp., Armonk, New York) predicting participants' reliance upon the advice of the care team member as a function of the advice giver's gender, experience, and role. We controlled for other factors (participants' own years in practice, gender, age, and practice type) to isolate the effect of advice giver attributes on participants' reliance on the treatment advice.

Results

Our results are summarized in Table 1, Table 2, and Figure 1. We note that some participant characteristics (such as gender and age) were significant predictors of advice taking in the scenario when the advice giver was a physician (Table 2). However, we focus our analyses here on the main effects and the interaction effects of our experimental manipulations on advice taking, controlling for participant characteristics that might explain additional variance but are not central to the focus of this study.

We found notably different effects of manipulated gender and experience on advice taking depending on whether the advice giver was a nurse or physician. Participants' reported reliance on advice from a nurse did not differ depending on the nurse's gender or experience (Table 1 and top panel of Figure 1). However, when advice was provided by a fellow physician, there was a significant gender \times experience interaction ($P = .01$; Table 2). This indicated 2 patterns consistent with gender bias in participants' advice taking, as shown in the bottom panel of Figure 1. First, when treatment advice was delivered by an inexperienced physician, participants reported relying significantly more on the advice of a man versus a woman, $F(1,61) = 4.24, P = .04$. Second, participants' reliance on the advice of the woman physician was a function of her experience, differing significantly across the low- and high-experience vignettes, $F(1,62) = 6.96, P = .01$, but their reliance on the advice of the man physician did not differ across his levels of experience, $F(1,60) = 0.21, P = .65$.

Discussion and Recommendations

Our findings are largely consistent with what is known about gender bias in the workplace.^{6,7,9,12} When the advice-giving physician in the vignette was a man, participants assumed his competence;

Table 1

General Linear Model Results Predicting Participants' Reliance on the Nurse's Advice as a Function of the Nurse's Manipulated Gender and Experience^a

Variables	Advice giver: Nurse	
	F statistic	P value (2-tailed)
Control variables: Participant characteristics		
Gender	.25	.62
Age	.62	.43
Years in practice	.81	.37
Practice type	.20	.65
Random variables: Vignette response		
Decision to intubate	2.54	.11
Confidence in decision	.49	.69
Independent variables: Nurse characteristics		
Gender (man vs woman)	.17	.68
Experience (15 years vs 2 years)	1.44	.23
Gender \times experience interaction	.03	.86

^aParticipants (283 anesthesiologists) read a brief online clinical vignette in which they were given a patient's history and vitals; they were then given treatment advice (intubate the patient) by a care team member (nurse or physician) within the vignette. Each participant received advice from 1 of the 8 possible advice givers, 4 of whom were nurses. Participants were asked whether they would intubate and to rate their level of reliance on the advice given by the care team member (see Figure 1). For the vignette, see Box 1.

his level of experience did little to increase or decrease their reliance on his advice. However, when the advice-giving physician was a woman, participants' perceptions of her competence (in the form of reliance on her advice) were

based on her level of experience. Our findings suggest that women physicians, relative to men, may encounter additional hurdles to performing their jobs, especially at early stages in their careers. These hurdles are rooted in the

Table 2

General Linear Model Results Predicting Participants' Reliance on the Physician's Advice as a Function of the Physician's Manipulated Gender and Experience^a

Variables	Advice giver: Physician	
	F statistic	P value (2-tailed)
Control variables: Participant characteristics		
Gender	4.49	.04
Age	5.39	.02
Years in practice	4.01	.04
Practice type	.68	.41
Random variables: Vignette response		
Decision to intubate	1.37	.24
Confidence in decision	1.05	.37
Independent variables: Physician characteristics		
Gender (man vs woman)	.02	.89
Experience (15 years vs 2 years)	2.70	.10
Gender \times experience interaction	6.33	.01

^aParticipants (283 anesthesiologists) read a brief online clinical vignette in which they were given a patient's history and vitals; they were then given treatment advice (intubate the patient) by a care team member (nurse or physician) within the vignette. Each participant received advice from 1 of the 8 possible advice givers, 4 of whom were physicians. Participants were asked whether they would intubate and to rate their level of reliance on the advice given by the care team member (see Figure 1). For the vignette, see Box 1.



Figure 1 Participating anesthesiologists' reported reliance on treatment advice from a care team member: a nurse (top panel) or a physician (bottom panel). In the study's online clinical decision-making vignette, each participant received advice to intubate from 1 of 8 possible advice givers (man or woman; physician or nurse; experienced or inexperienced; see Box 1 for the vignette). The data shown here are from participants' responses ($n = 283$) to the post-vignette question, "When you made your decision, to what degree did you rely upon the advice of the other care team member?" Responses used a 5-point scale (0 = not at all to 4 = very much).

^aSignificant pairwise comparison, $P = .04$.

^bSignificant pairwise comparison, $P = .01$.

psychological biases of others, rather than in the objective features of a case or a treatment setting.

Gender bias in assumed competence is a phenomenon that emerges early in physicians' careers.¹² For example, recent research suggests that medical faculty tend to give more autonomy to men versus women thoracic surgery residents based solely on residents' gender.¹³ Our data suggest that even when women and men engage in the same autonomous act (offering advice to a colleague), that act may carry less weight when initiated by a woman versus a man. In light of known gender disparities in medicine with respect to salary, advancement, and exposure to sexual harassment and microaggressions—all of which disadvantage women relative to men¹⁴—taking action to address subtle forms of gender bias, like those observed in this study, is a critical step toward achieving a health care system in which talent and

expertise, rather than social category, determine the value of one's perspective.

It is noteworthy that gender bias in advice taking was not observed when treatment advice was offered by a nurse. Figure 1 suggests that advice taking was lower overall when advice was offered by a nurse versus a physician, regardless of the nurse's gender or experience. Known status differences within medicine for physicians and nurses¹⁵ suggest that this result could be explained in terms of a general tendency to discount advice offered by lower-status actors. If so, gender bias may not have emerged because the nurse advice giver was already viewed as lower status by the participating physicians.

Additional research is needed to clarify and extend other aspects of our findings. For example, our results show that advice taking varied as a function of the characteristics of the participants

themselves (see Table 2), suggesting that our primary findings might be more or less pronounced for women versus men physicians or for more versus less experienced physicians. Research studies with fewer experimental conditions and/or larger sample sizes would offer a better opportunity to understand individual-level moderators of these findings. Further, extending this research into other specialties, particularly those with a more balanced gender composition among practicing physicians,¹⁶ would speak to how the gender composition of a medical specialty contributes to gender bias in advice taking. Finally, future research should examine biases in advice taking as a function of other socially relevant categories, such as the race/ethnicity or age of the advice giver.

At a broad level, this exploratory study reflects growing efforts to apply behavioral science to understand the human aspects of medical decision making. In the last 50 years, similar efforts have refined basic assumptions about human decision making in fields including economics, finance, and public health, and recent calls from within medicine have acknowledged a need to understand physician decision making through this lens.^{17,18} Our findings show that cultural stereotypes may shape physicians' information use and decision-making processes (and hinder collaboration), even in contexts that appear to have little to do with social category membership. In our study, the decision makers were experts in their field, which further underlines the insidiousness of such stereotypes.

Practicing physicians should be aware of their own susceptibility to gender bias and its potential detrimental effects on care decisions. Fostering formal awareness—for example, by addressing gender bias in clinical exams and residency programs—is likely a critical step toward identifying solutions.¹⁹ However, awareness alone is insufficient. Indeed, recent organizational behavior research suggests interventions such as mandatory unconscious bias training, a mainstay in many organizational settings, do little to eradicate bias based on social attributes such as gender or race.²⁰ Moreover, a recent study of microaggressions in medicine suggests that low-level instances of gender-based discrimination are less likely

to be detected by men versus women observers,²¹ suggesting that the likeliest perpetrators of these acts may also be the least able to recognize when they are engaged in them.

A more useful approach would be to adopt formal systems for addressing specific problems.¹⁹ In the case of our research context, adopting policies and practices that encourage equal attention to advice, regardless of the source, could help ensure that advice taking is a function of information quality rather than the irrelevant attributes of the advice giver. Such policies and practices may help surface and implement diverse expert perspectives in collaborative medical decision making, promoting better and more effective patient care.

Funding/Support: Internal funding reported.

Other disclosures: None reported.

Ethical approval: Johns Hopkins Homewood Institution Review Board Protocol HIRB00007911, approved October 20, 2018.

Disclaimers: The first author is an employee of the U.S. Government. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

Previous presentations: Johns Hopkins University departmental brown bag events.

E.G. Helzer is associate professor, Naval Postgraduate School, Monterey, California.

C.G. Myers is assistant professor, Johns Hopkins Carey Business School and Johns Hopkins School of Medicine, Baltimore, Maryland.

C. Fahim is a scientist, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada.

K.M. Sutcliffe is professor, Johns Hopkins Carey Business School and Johns Hopkins School of Medicine, Baltimore, Maryland.

J.H. Abernathy is associate professor of anesthesiology and critical care medicine, Johns Hopkins School of Medicine, Baltimore, Maryland.

References

- 1 Thomas DB, Newman-Toker DE. Diagnosis is a team sport—Partnering with allied health professionals to reduce diagnostic errors: A case study on the role of a vestibular therapist in diagnosing dizziness. *Diagnosis*. 2016;3:49–59.
- 2 Marewski JN, Gigerenzer G. Heuristic decision making in medicine. *Dialogues Clin Neurosci*. 2012;14:77–89.
- 3 Henrich J, Boyd R, Young P, et al. What is the role of culture in bounded rationality? In: Gigerenzer G, Selten R, eds. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press; 2002:343–359.
- 4 Hilton JL, von Hippel W. Stereotypes. *Annu Rev Psychol*. 1996;47:237–271.
- 5 Nisbett RE, Wilson TD. Telling more than we can know: Verbal reports on mental processes. *Psychol Rev*. 1977;84:231–259.
- 6 Thomas-Hunt MC, Phillips KW. When what you know is not enough: Expertise and gender dynamics in task groups. *Pers Soc Psychol Bull*. 2004;30:1585–1598.
- 7 Joshi A. By whom and when is women's expertise recognized? The interactive effects of gender and education in science and engineering teams. *Admin Sci Quart*. 2014;59:202–239.
- 8 Botelho TL, Abraham M. Pursuing quality: How search costs and uncertainty magnify gender-based double standards in a multistage evaluation process. *Admin Sci Quart*. 2017;62:698–730.
- 9 Hengle E. Publishing while female: Are women held to higher standards? Evidence from peer review. http://www.erinhengel.com/research/publishing_female.pdf. Published March 2020. Accessed June 30, 2020.
- 10 Rochon PA, Davidoff F, Levinson W. Women in academic medicine leadership: Has anything changed in 25 years? *Acad Med*. 2016;91:1053–1056.
- 11 Fnais N, Soobiah C, Chen MH, et al. Harassment and discrimination in medical training: A systematic review and meta-analysis. *Acad Med*. 2014;89:817–827.
- 12 Myers CG, Sutcliffe KM. How discrimination against female doctors hurts patients. *Harvard Business Review*. <https://hbr.org/2018/08/how-discrimination-against-female-doctors-hurts-patients>. Published August 30, 2018. Accessed June 30, 2020.
- 13 Meyerson SL, Sternbach JM, Zwischenberger JB, Bender EM. The effect of gender on resident autonomy in the operating room. *J Surg Educ*. 2017;74:e111–e118.
- 14 Raj A, Kumra T, Darmstadt GL, Freund KM. Achieving gender and social equality: More than gender parity is needed. *Acad Med*. 2019;94:1658–1664.
- 15 Thomas CD, DiSabatino L, Rojas F. The effects of education and clinical specialization on nurses' status affirmation by physicians: A quantitative analysis. *J Interprof Care*. 2019;33:252–263.
- 16 Association of American Medical Colleges. Table 2: Distribution of residents by specialty, 2005 compared to 2015. <https://www.aamc.org/system/files/reports/1/2015table2.pdf>. Published 2016. Accessed July 16, 2020.
- 17 Avorn J. The psychology of clinical decision making—Implications for medication use. *N Engl J Med*. 2018;378:689–691.
- 18 Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. *Annu Rev Public Health*. 2010;31:399–418.
- 19 McKinley SK, Wang LJ, Gartland RM, et al; Massachusetts General Hospital Gender Equity Task Force. "Yes, I'm the doctor": One department's approach to assessing and addressing gender-based discrimination in the modern medical training era. *Acad Med*. 2019;94:1691–1698.
- 20 Chang EH, Milkman KL, Gromet DM, et al. The mixed effects of online diversity training. *Proc Natl Acad Sci U S A*. 2019;116:7778–7783.
- 21 Periyakoil VS, Chaudron L, Hill EV, Pellegrini V, Neri E, Kraemer HC. Common types of gender-based microaggressions in medicine. *Acad Med*. 2020;95:450–457.