

# Recognizing and Mitigating Gender Bias in Medical Teaching Assessments

Jessica C. Babal, MD  
Sarah Webber, MD  
Carrie L. Nacht, MPH  
Kirstin A.M. Nackers, MD

Kristin Tiedt, MD  
Ann Allen, MD  
Brittany J. Allen, MD  
Michelle M. Kelly, MD, MS

## Introduction

Gender bias in graduate medical education (GME) is well-documented.<sup>1-6</sup> Research and mitigation strategies are largely directed at gender bias within resident performance assessments. However, evidence suggests that gender bias also appears in faculty teaching assessments<sup>2,3</sup> and that long-standing gender inequities in academic medicine may persist in part because of the “culmination of countless ‘small’ differences” in how faculty are assessed.<sup>6</sup> Therefore, to mitigate gender bias in GME, we must recognize bias throughout the educational hierarchy and modify structures that facilitate its impact. Here, we draw attention to gender bias in GME teaching assessments and propose several bias mitigation strategies.

## Intent Versus Impact

Teaching assessments are fundamental to medical education and ideally facilitate faculty professional development. Assessments may bring attention to outstanding teaching to reward and problematic teaching to address.<sup>7</sup> However, despite intent or desire for objectivity, teaching assessments may harbor biases and may speak more to a faculty person’s ability to adhere to normative or expected behavior for gender rather than to their teaching skills.<sup>8</sup>

## Gender Bias and Expectations

Gender biases are assumptions or perceptions one holds about gender. Gender biases may be *implicit* (implied, intuited) or *explicit* (identified, expressed).<sup>1</sup> Biases enable mental shortcuts and may become deeply ingrained. Consequently, persons of all genders hold gender biases. Many gender biases are unconscious and may substantially differ from self-identified beliefs about gender. Biases persist despite generational progress in gender equality.<sup>1</sup>

Gender biases are informed by long-standing cultural expectations for how individuals should act.

Traditional expectations suggest that gender is *binary* (man/woman) and that gender expression of *masculinity* or *femininity* should align with social expectations for sex assigned at birth.<sup>9</sup> Traditionally, masculinity evokes expectations of assertiveness, leadership, and technical skills, while femininity evokes expectations of caregiving, relationship building, and teamwork.<sup>1</sup>

When an individual’s gender expression does not align with expectations of masculinity or femininity, they may face backlash. For example, cisgender women (whose gender aligns with sex designated at birth) who demonstrate stereotypically masculine traits (eg, assertiveness) commonly face criticism, particularly in specialties with low representation of women.<sup>1,3,10</sup> Similarly, men may face social reprisal for displaying stereotypically feminine traits (eg, emotional expressiveness).<sup>1</sup> Although cisgender men physicians in women-predominant specialties still outperform women counterparts in promotion and pay,<sup>11-13</sup> men in these specialties may face interpersonal bias for working in specialties viewed as feminine.<sup>14</sup> Transgender and gender diverse (TGD) individuals experience additional scrutiny<sup>9</sup> that remains underexplored to date.

## Intersectionality

Gender bias cannot be disentangled from other social biases. When we encounter others, we do not simply recognize gender, but rather the intersection of identities, including race, ethnicity, religion, sexuality, disability, and body type, all of which may elicit additional biases.<sup>9</sup> In particular, cisgender women, cisgender men, and TGD individuals whom identify as Black, Indigenous, (and) People of Color (BIPOC), or other historically marginalized identities face immense bias and discrimination in and outside of medicine.<sup>15,16</sup>

## Gender Bias and Implications for Success

The FIGURE shows 3 hypothetical teaching assessments demonstrating gender bias. In the examples, all 3

DOI: <http://dx.doi.org/10.4300/JGME-D-21-00774.1>

<p><b>Dr. Harris</b></p> <ul style="list-style-type: none"> <li>• Cisgender White man</li> <li>• Mid-career internist</li> </ul>	<p>“He is knowledgeable and a clear leader on the wards. Dedicated physician. A role model in challenging situations. Teaching points were high yield.”</p>
<p><b>Dr. Robinson</b></p> <ul style="list-style-type: none"> <li>• Cisgender Black woman</li> <li>• Mid-career internist</li> </ul>	<p>“Committed physician. Patients like her. She is extremely knowledgeable. However, she is less approachable than other attendings.”</p>
<p><b>Dr. Gray</b></p> <ul style="list-style-type: none"> <li>• Transgender Indigenous man</li> <li>• Mid-career internist</li> </ul>	<p>“Committed to patients and learners. Had some good teaching points.”</p>

FIGURE

Hypothetical Examples of Gender Bias in Faculty Teaching Assessments

internists were praised for being committed physicians. However, the resident primarily emphasized the cisgender man’s *agency* (leadership) and the cisgender woman’s *communal* (relationships).<sup>1</sup> Additionally, *doubt*-inducing language (“however”) directly followed praise for the cisgender woman’s agentic quality (“knowledgeable”), introducing uncertainty about performance success.<sup>17</sup> For the transgender internist, we see no explicit discrimination language; however, the brevity and relatively lukewarm response may indicate uncertainty about approaches to assessing a person who challenges the gender binary.<sup>9,18</sup>

Traditional views of leadership in medicine treat masculine traits as primary markers of success. Therefore, traditionally feminine descriptors in teaching assessments may elicit unconscious assumptions that a person has lesser performance potential.<sup>17</sup> This may not only have implications for professional advancement but may also impact professional identity formation and self-evaluation (eg, *imposter syndrome*).<sup>7,10</sup> Moreover, gender-based microaggressions and discrimination in assessments may intensify harm by reinforcing *stereotype threats*—concerns about conforming to negative stereotypes about one’s social group—negatively affecting performance and perpetuating equity gaps.<sup>19</sup>

## How Can GME Mitigate Gender Bias in Teaching Assessments?

### Strategy 1: Implement Individual Behavior Change by Using Language That Treats Gender as a Spectrum

By continuing to describe gender as binary, instead of the continuum that it is, we perpetuate gender bias.<sup>9</sup> Seeking to understand inequities *across the gender spectrum* instead of inequities *between men and women*, serves to disrupt the current frameworks many of us have built into our mindsets.<sup>9</sup> Examples include using correct terminology to describe patient self-identified gender in case presentations (eg, cisgender male). Similarly, systematically treating gender as a spectrum influenced by identity intersectionality will bring attention to the impact of gender bias on a broader range of identities.

### Strategy 2: Broaden Institutional Understanding of Problematic Assessment Tactics

Several assessment approaches perpetuate gender bias by tapping into intuitive parts of the evaluator brain where potential biases reside. Nonspecific and trait-based questions are particularly problematic (eg, What are this teacher’s strengths?)<sup>10</sup> Such questions encourage residents to rely on intuition about *who* makes a good teacher and *how* that teacher should conduct themselves based on gender expectations.<sup>10</sup>

**TABLE**  
Example Behavior-Based Teaching Assessment Questions

Teaching Skill <sup>20</sup>	Example Behavior-Based Question <sup>a</sup> During the rotation, how often did the attending...
Teaches medical or surgical condition management	... use teaching strategies that advanced your understanding of medical (surgical) condition management?
Supports procedural skill development	... provide clear explanations of proper procedural techniques?
Grants appropriate autonomy	... allow for your autonomy in medical decision-making to an extent that was appropriate for your skill level?
Provides patient care or procedural support when needed	... provide you with patient care assistance when needed (eg, when having difficulty in communication with a patient)?
Provides effective feedback	... provide actionable feedback on improving your patient care plans?
Ensures supportive learning environment	... demonstrate patience when you asked questions about treatment plans?
Demonstrates professionalism	... arrive on time for rounds?

<sup>a</sup> Response options: More than 75% of the time; 51%–75% of the time; 25%–49% of the time; less than 25% of the time; not applicable.<sup>21</sup>

Note: Although free-response questions are potentially subject to greater bias than frequency-based questions due to their open-endedness,<sup>10</sup> each question could be developed to elicit free-text responses by asking what behaviors the attending used effectively to demonstrate each teaching skill and to elicit recommendations for teaching skill development. For example, during this rotation, what strategies did the attending use that advanced your understanding of medical condition management?

Similarly, bias likelihood increases after brief teaching interactions because these interactions provide fewer concrete examples of teaching behaviors.<sup>10</sup> Bias likelihood also increases when significant time passes following a teaching interaction, because memories of specific teaching behaviors degrade over time, encouraging reliance on intuition.<sup>10</sup> In contrast, timely, specific, and behavior-based questions (TABLE) following more extensive interactions encourage residents to draw on concrete examples of teaching behaviors. Educational workshops<sup>22</sup> may broaden institutional understanding of assessment tactics that may perpetuate bias.

### Strategy 3: Develop Institutional Procedures to Evaluate Assessments

Creating institutional procedures that encourage best assessment practices could systematically reduce bias.<sup>23</sup> This might include developing a stakeholder workgroup that evaluates performance assessments and advises institutional leaders on assessment development. Specifically, workgroups might evaluate the validity, reliability, and utility of assessments,<sup>24</sup> asking:

- Does this question more effectively assess the teacher's skills or adherence to gender expectations?
- Would this question be answered similarly for a cisgender woman, a cisgender man, and a TGD individual?
- What purpose does this question serve? In pursuit of its intended aim, might this question perpetuate gender inequities?

If an assessment question presents gender bias concerns, revision should be considered. Because even optimized questions may not eliminate bias, institutions might use an electronic prompt at the beginning of teaching assessments reminding residents about gender bias risk.<sup>25</sup> All changes should then be evaluated<sup>26</sup> to ensure that efforts intended to support gender equity do not overlook issues relating to intersectionality identities or create new equity challenges.<sup>27</sup>

### Strategy 4: Address Gender Discrimination and the Leaky Pipeline Within and Across Institutions

Mitigating gender bias will require more than assessment template changes. Implementing anti-sexism and anti-discrimination programs and policies may challenge gender expectations and inequities perpetuated over generations.<sup>16</sup> Moreover, repairing the leaky pipeline that has particularly excluded BIPOC women and men and TGD persons in academic medicine may reduce gender bias.<sup>28</sup> Increasing representation of BIPOC individuals and TGD persons within leadership may help shift cultural perceptions about gender/identity and performance capabilities<sup>9,10</sup> and contribute to ongoing progress in developing equitable systems.

### Conclusion

Gender bias in GME is a pervasive influencer of gender inequities. The influence of gender on teaching assessments warrants further attention. Concerted action that aims to recognize and address gender bias in teaching assessments may be a starting point in reducing inequities.

## References

1. Carnes M, Bartels CM, Isaac C, Kaatz A, Kolehmainen C. Why is John more likely to become department chair than Jennifer? *Trans Am Clin Climatol Assoc.* 2015;126:197–214.
2. Heath JK, Weissman GE, Clancy CB, Shou HC, Farrar JT, Dine CJ. Assessment of gender-based linguistic differences in physician trainee evaluations of medical faculty using automated text mining. *JAMA Netw Open.* 2019;2(5):e193520. doi:10.1001/jamanetworkopen.2019.3520
3. Fassiotto M, Li J, Maldonado Y, Kothary N. Female surgeons as counter stereotype: the impact of gender perceptions on trainee evaluations of physician faculty. *J Surg Educ.* 2018;75(5):1140–1148. doi:10.1016/j.jsurg.2018.01.011
4. 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
5. Hemphill ME, Maher Z, Ross HM. Addressing gender-related implicit bias in surgical resident physician education: a set of guidelines. *J Surg Educ.* 2020;77(3):491–494. doi:10.1016/j.jsurg.2019.12.014
6. Sheffield V, Hartley S, Stansfield RB, et al. Gendered expectations: the impact of gender, evaluation language, and clinical setting on resident trainee assessment of faculty performance [published online ahead of print August 17, 2021]. *J Gen Intern Med.* doi:10.1007/s11606-021-07093-w
7. Vaughan B. Clinical educator self-efficacy, self-evaluation and its relationship with student evaluations of clinical teaching. *BMC Med Educ.* 2020;20(1):347. doi:10.1186/s12909-020-02278-z
8. Templeton K. Impact of gender on teaching evaluations of faculty: another example of unconscious bias? *J Womens Health (Larchmt).* My 2016;25(5):420–421. doi:10.1089/jwh.2016.29008.kte
9. Morgenroth T, Ryan MK. The effects of gender trouble: an integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspect Psychol Sci.* 2021;16(6):1113–1142. doi:10.1177/1745691620902442
10. Heilman ME. Gender stereotypes and workplace bias. *Res Org Behav.* 2012;32:113–135. doi:10.1016/j.riob.2012.11.003
11. Rotbart HA, McMillen D, Taussig H, Daniels SR. Assessing gender equity in a large academic department of pediatrics. *Acad Med.* 2012;87(1):98–104. doi:10.1097/ACM.0b013e31823be028
12. Hoffer L, Hacker MR, Dodge LE, Ricciotti HA. Subspecialty and gender of obstetrics and gynecology faculty in department-based leadership roles. *Obstet Gynecol.* 2015;125(2):471–476. doi:10.1097/AOG.0000000000000628
13. Eichelberger KY. Equal pay for equal work in academic obstetrics and gynecology. *Obstet Gynecol.* 2018;131(2):224–226. doi:10.1097/AOG.0000000000002420
14. Long TR, Arendt KW, Elliott BA, Rose SH. Gender and graduate medical education: is obstetrics and gynecology a justifiable outlier? *Mayo Clin Proc.* 2018;93(6):684–688. doi:10.1016/j.mayocp.2018.02.025
15. Avakame EF, October TW, Dixon GL. Antiracism in academic medicine: fixing the leak in the pipeline of Black physicians. *ATS Sch.* 2021;2(2):193–201. doi:10.34197/ats-scholar.2020-0133PS
16. Kvach EJ, Weinand J, O'Connell R. Experiences of transgender and nonbinary physicians during medical residency program application. *J Grad Med Educ.* 2021;13(2):201–205. doi:10.4300/JGME-D-20-00384.1
17. Trix F, Psenka C. Exploring the color of glass: letters of recommendation for female and male medical faculty. *Discourse & Society.* 2016;14(2):191–220. doi:10.1177/0957926503014002277
18. Poteat T, German D, Kerrigan D. Managing uncertainty: a grounded theory of stigma in transgender health care encounters. *Soc Sci Med.* 2013;84:22–29. doi:10.1016/j.socscimed.2013.02.019
19. Schmader T, Johns M, Forbes C. An integrated process model of stereotype threat effects on performance. *Psychol Rev.* 2008;115(2):336–356. doi:10.1037/0033-295X.115.2.336
20. Pourmand K, Sewell JL, Shah BJ. What makes a good endoscopic teacher: a qualitative analysis. *J Surg Educ.* 2018;75(5):1195–1199. doi:10.1016/j.jsurg.2018.02.016
21. Sullivan GM, Artino AR Jr. How to create a bad survey instrument. *J Grad Med Educ.* 2017;9(4):411–415. doi:10.4300/JGME-D-17-00375.1
22. Harrison-Bernard LM, Augustus-Wallace AC, Souza-Smith FM, Tsien F, Casey GP, Gunaldo TP. Knowledge gains in a professional development workshop on diversity, equity, inclusion, and implicit bias in academia. *Adv Physiol Educ.* 2020;44(3):286–294. doi:10.1152/advan.00164.2019
23. American Association of University Professors. Statement on Teaching Evaluation. Accessed November 1, 2021. <https://www.aaup.org/report/statement-teaching-evaluation>
24. Accreditation Council for Graduate Medical Education. Holmboe ES, Lobst WF. Assessment Guidebook. Accessed November 1, 2021. <https://www.acgme.org/globalassets/PDFs/Milestones/Guidebooks/AssessmentGuidebook.pdf>
25. Peterson DAM, Biederman LA, Andersen D, Ditonto TM, Roe K. Mitigating gender bias in student

- evaluations of teaching. *PloS One*. 2019;14(5):e0216241. doi:10.1371/journal.pone.0216241
26. Erdahl LM, Chandrabose RK, Pitt SC, Radford DM, Strong SA, Silver JK. A call for professionalism: addressing gender bias in surgical training. *J Surg Educ*. 2020;77(4):718–719. doi:10.1016/j.jsurg.2020.02.004
  27. Tzanakou C. Unintended consequences of gender-equality plans. *Nature*. 2019;570(7761):277. doi:10.1038/d41586-019-01904-1
  28. Lett E, Orji WU, Sebro R. Declining racial and ethnic representation in clinical academic medicine: a longitudinal study of 16 US medical specialties. *PloS One*. 2018;13(11):e0207274. doi:10.1371/journal.pone.0207274



All authors are with the University of Wisconsin School of Medicine and Public Health. **Jessica C. Babal, MD**, is Assistant Professor (CHS), Department of Pediatrics, Division of General

Pediatrics and Adolescent Medicine; **Sarah Webber, MD**, is Assistant Professor (CHS), Department of Pediatrics, Division of Hospital Medicine, and Department of Pediatrics Well-Being Director; **Carrie L. Nacht, MPH**, is Research Specialist, Department of Pediatrics, Division of Hospital Medicine; **Kirstin A.M. Nackers, MD**, is Associate Professor (CHS), Department of Pediatrics, Division of Hospital Medicine, and Director for Medical Student Education; **Kristin Tiedt, MD**, is Assistant Professor (CHS), Department of Pediatrics, Division of Hospital Medicine; **Ann Allen, MD**, is Assistant Professor (CHS), Department of Pediatrics, Division of Hospital Medicine, Medical Director of Community Pediatric Hospital Medicine, and Program Director, Pediatric Hospital Medicine Fellowship; **Brittany J. Allen, MD**, is Associate Professor (CHS), Department of Pediatrics, Division of General Pediatrics and Adolescent Medicine, and Co-Medical Director, Pediatric and Transgender Health (PATH) Clinic, American Family Children's Hospital; and **Michelle M. Kelly, MD, MS**, is Associate Professor (CHS), Department of Pediatrics, Division of Hospital Medicine.

The authors would like to thank Megan Moreno, MD, for her review of an earlier version of this manuscript and to Christine Richards for support in formatting and preparing this manuscript.

Corresponding author: Jessica C. Babal, MD, University of Wisconsin School of Medicine and Public Health, babal@wisc.edu