



# Reducing Bias in Clerkship Evaluations

Emily Nguyen, MD, FAAP; Vydia Permashwar, MBBS FAAP  
 Carilion Childrens Hospital, Department of Pediatrics

Authors have no financial relationships to disclose.



## Background

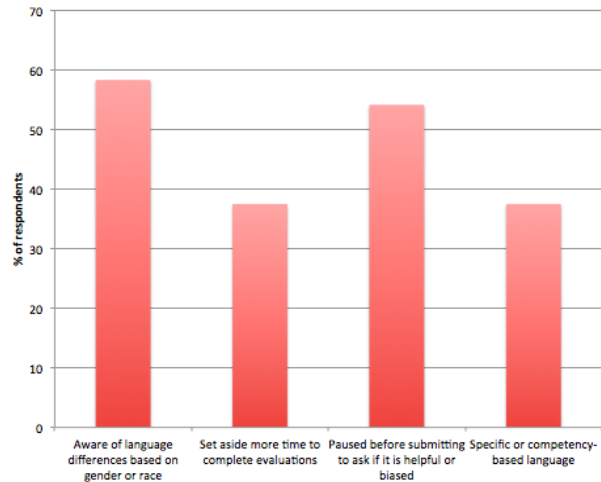
Previous studies have demonstrated that only 44% of medical students believed clerkship grading was fair, and 67% reported concerns that clerkship grades and feedback were not based solely on performance<sup>1</sup>. White students are more likely to be described with standout words such as “exceptional,” compared to Blacks, Hispanics, and Asians, even after controlling for USMLE scores; women applicants were more likely to be described with words relating to compassion<sup>2</sup>. Smaller differences in evaluations may lead to an “amplification cascade” in grades with URM students less likely to be nominated for prestigious societies such as AOA and match into competitive residency programs<sup>3</sup>. Faculty may not be aware of differences in language used to evaluate medical students or bias regarding gender or race/ethnicity.

## Methods

We developed and presented a 30-minute Faculty Development session at Pediatric Grand Rounds targeting faculty and residents involved in medical education. The session included a literature review on bias in clerkship evaluations based on gender and race. An interactive exercise was utilized to practice recognizing and addressing biased language in evaluation; examples and strategies were provided to help counteract bias.

Evaluations were collected and analyzed following the session with an additional follow-up survey distributed 3.5 months later to assess the self-reported longer-term impact on behavior.

## Results



- 43 Carilion and non-Carilion faculty and residents attended the session. 42% of attendees completed the initial survey, with 94% rating the session as “excellent” or “very good,” and 94% stating it was relevant to their needs as faculty.
- 56% of attendees completed follow-up surveys.
- 42% of these had not before received training on how to evaluate learners. The most implemented change was pausing to ask if feedback was helpful or biased (54%) and greater awareness in gender or race based language differences (58%).
- Qualitatively, participants reported increased competency in identifying and minimizing bias in evaluation.

## Conclusion

We provided education to pediatric faculty and residents on the role of bias in written evaluations and strategies to counteract it, with overall favorable evaluations on both content and relevance. Most participants reported they became more aware of differences in language used to describe learners based on gender or race. Based on survey results obtained 3.5 months post-session, participants reported having made several changes to make student evaluations more helpful and less biased, suggesting brief education may have longer-term impact on behavior.

## Future directions

Additional faculty development sessions to increase assessors’ skills, quality of feedback and extent of bias may help make clerkship evaluations more objective or competency-based and may be extrapolated to other learners/evaluator dyads.

Future directions include evaluating the longer-term impact of these sessions in effecting sustained change in faculty behavior. Expansion involving faculty from other core rotations will have a wider impact across the medical school.

## References

1. Bullock, et al. “Healing a Broken Clerkship Grading System.” AAMC, Feb. 2020
2. Ross DA, et al. (2017) Differences in words used to describe racial and gender groups in Medical Student Performance Evaluations. PLoS ONE 12(8): e0181659
3. Teherani A, et al. How Small Differences in Assessed Clinical Performance Amplify to Large Differences in Grades and Awards: A Cascade With Serious Consequences for Students Underrepresented in Medicine. Acad Med. 2018 Sep;93(9):1286-1292.