Countering Bias in Assessment

Adelaide H. McClintock, MD (@McClintockMD)
Tyra Fainstad, MD

Joshua Jauregui, MD (@joshuajauregui)
Lalena M. Yarris, MD, MCR (@lainieyarris)

The Challenge

Assessments can provide meaningful performance data from expert observers, but this information is prone to harmful bias. Such assessment bias disproportionately affects trainees who do not resemble or share identities with those doing the assessment.[1] While hundreds of cognitive biases exist, some have particularly pernicious influences when building and sustaining diversity in medicine, with subtle differences in individual performance ratings systematically perpetuating the exclusion of marginalized groups.[1]

What Is Known

Bias affects our perceptions of another's knowledge, ability, professionalism, and readiness for independent practice.[2] Over repeated assessments, these biases can result in an amplification cascade,[1] a phenomenon in which small differences in assessed performance lead to larger differences in grades and selection for awards, favoring well-represented individuals and hindering underrepresented in medicine (UiM) trainees in achieving training success.[3] Eliminating harmful effects of cognitive biases requires a multilevel response. Recognizing that bias and cognitive error cannot be "trained out" of individuals,[4] systems can be put in place to create more equitable assessment. We must start by critically evaluating the tools, processes, and outcomes of our existing assessments.[3] Equitable assessments are criterion-based, encompass multiple dimensions of optimal patient care, and take place in the context of a longitudinal relationship.[5] Assessment should be reoriented from a deficit-focused lens toward a growth-focused model that uses goal setting, differentiated assessments, direct observation, and frequent feedback, both positive and critical.[5]

How You Can Start TODAY

1. **Require anti-bias training for all supervising clinicians.** Training should facilitate faculty recognition of their own biases, teach faculty to recognize biased language in their own narrative assessments, and provide skills to counter biased thinking.[2] Create an awareness of how bias influences trainee assessment and the significance of the amplification cascade.

2. **Establish clear criteria for competency-based assessments.** Many assessment forms have already moved from normative to criterion-based assessment

---

> **RIP OUT ACTION ITEMS**
>
> 1. Educate faculty about adverse effects of assessment bias; provide training to recognize and address this bias.
> 2. Reconsider the purpose of assessment. Focus on learning and professional growth for attainment of criterion-based competencies prior to independent practice.
> 3. (Re)build assessment systems that are criterion-based, reward multiple dimensions of patient care, and promote trainee growth, all in the context of longitudinal relationships.
> 4. Invite trainees to meaningfully participate in individual goal setting, curriculum changes, instrument design, and the processes for defining and measuring outcomes.

(eg, pass/fail) to address inequity in advancement, including the United States Medical Licensing Examination Step 1 examination and medical school clerkship grades. Assessments of physician readiness should similarly use clear criteria for "met" or "not met" to assess readiness for independent practice.

3. **Prioritize assessment for learning.** Formative assessment should be incorporated into feedback practices that facilitate learning and progression toward safe, independent practice. Once trainees have met appropriate skill and safety metrics, assessments should shift to focus on growth, with trainee-adviser co-constructed learning goals.

4. **Name, reframe, and check-in.**[6] Build the expectation that assessment and feedback are a bidirectional dialogue. Before providing feedback, assessors should describe their expectations and standards. Assessors should also name the presence of bias directly to trainees. Use language that acknowledges bias, reflects subjectivity of human judgement, and focuses on observed behaviors. For example, rather than "you are a good communicator," reframe to: "*I think* the patient understood your directions since they were able to repeat them back to you." Check-in with trainees to assess whether feedback seems relevant, true to them, and action oriented.

5. **Reframe trainee response.** Trainees may seem defensive regarding feedback. Assessors can reframe this as trainee perspectives on situational and contextual information. For example, a trainee may reply to your feedback about improving a specific communication skill with information about how they built rapport that occurred prior to coming in the room.

What You Can Do LONG TERM

1. **Examine your individual and program values.** Use of narrowly defined, favored norms can marginalize UiM trainees. Review your work-based assessment tools and processes to discern if certain skills or traits are routinely favored and rewarded. Are growth orientation, reflective practice, or humility assessed? Where and how are these observed and measured? Does your program have multiple strategies for trainees to demonstrate competency? Seek input from trainees and patient advocacy representatives to broaden and appropriately reward multiple definitions of success and growth.

2. **Provide data transparency at the program level.** Analyze and share program-specific assessment data, looking for differences across groups (eg, gender, race, country of origin). Scrutinize current assessment data with checkpoints at each level of data aggregation and decision-making (work-based assessment, competency decisions, entrustment, and advancement). Seek to understand and explain how work-based assessment forms, narratives, or aggregate processes might be contributing to inequities. Develop standardized processes to remedy weak points.

3. **Build accountability at every level of the program.** Establish at least annual reviews of assessment data at the individual faculty level and incorporate them into the Annual Program Evaluation. Data reviews should be monitored by designated institutional officials and be reported to key stakeholders (eg, sponsoring institution, Accreditation Council for Graduate Medical Education). Examine narrative evaluations for gendered or other biased language using keyword searches (eg, wonderful, fabulous, good, pleasant, open, nice) or natural language processing.

4. **Define and track metrics of success.** Once specific sources of inequity are identified, downstream events such as advancement, fellowship match, job selection, and career advancement can then be followed to ensure progress toward opportunity equity.

References and Resources for Further Reading

1. Teherani A, Hauer KE, Fernandez A, King TE Jr, Lucey C. 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;93(9):1286–1292. doi:10.1097/ACM.0000000000002323
2. Ogunyemi D. Defeating unconscious bias: the role of a structured, reflective, and interactive workshop. *J Grad Med Educ*. 2021;13(2):189–194. doi:10.4300/JGME-D-20-00722.1
3. Lucey CR, Hauer KE, Boatright D, Fernandez A. Medical education's wicked problem: achieving equity in assessment for medical learners. *Acad Med*. 2020;95(suppl 12):98–108. doi:10.1097/ACM.0000000000003717
4. Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the "black box" differently: assessor cognition from three research perspectives. *Med Educ*. 2014;48(11):1055–1068. doi:10.1111/medu.12546
5. Teherani A, Perez S, Muller-Juge V, Lupton K, Hauer KE. A narrative study of equity in clinical assessment through the antideficit lens. *Acad Med*. 2020;95(suppl 12):121–130. doi:10.1097/ACM.0000000000003690
6. Fainstad TL, McClintock AH, Yarris LM. Bias in assessment: name, reframe, and check in [published online ahead of print March 30, 2021]. *Clin Teach*. doi:10.1111/tct.13351

Adelaide H. McClintock, MD, is Assistant Professor of Medicine, University of Washington; **Tyra Fainstad, MD,** is Visiting Associate Professor of Medicine, University of Colorado; **Joshua Jauregui, MD,** is Assistant Professor of Emergency Medicine, University of Washington; and **Lalena M. Yarris, MD, MCR,** is Professor of Emergency Medicine and Vice Chair for Faculty Development, Oregon Health & Science University, and Deputy Editor, *Journal of Graduate Medical Education*.

Corresponding author: Adelaide H. McClintock, MD, University of Washington, ahearst@uw.edu, Twitter @McClintockMD