Mitigating Bias in Assessment and Evaluation

Tomoko Sairenji, MD, MS
Jeanne Cawse-Lucas, MD
Genya Shimkin, MPH
Toby Keys, MPH
Susan Merel, MD
Disclosure

None
Disclaimers

• We are human and have our own biases
• We will discuss gender bias and bias against under-represented minorities in medicine (URM) – other forms of bias exist
• All of us here are doing the best we can . . . and make mistakes!
Objectives

1. Describe the impact of bias in assessment in evaluations
2. Identify language commonly used in written comments that reflect racial and gender bias
3. Use competency-based frameworks (e.g., EPAs and milestones) in writing evaluations
4. Create a personal plan to mitigate bias when evaluating learners
5. Plan changes at your institution to reduce biased assessments and evaluations
The literature
-Bias in assessment-
Racial disparities exist in clinical grading

• Examined clerkship grades and MSPE summary words
• Student categories: White, URM, non-URM minority
• In 4 out of 6 clerkships, grading disparities (p<0.05) favored White students over URM and non-URM minority students
• Non-URM minority students received lower MSPE summary words (p=0.001)
• Single institution study: UWSOM

Low et al. Teaching and Learning in Medicine. 31 (5) 2019
Bias is reflected in words used in assessment and evaluation

- Letters of Recommendation
- Clerkship evaluations
- MSPE
Letters of Recommendation

Men = **Agentic**
Decisive
Ambitious
Strong
Assertive
Logical
Independent

Women = **Communal**
Nurturing
Supportive
Caring
Kind
Emotional
Dependent

Men = **Standout**
Exceptional
Best
Outstanding
Superb

Women = **Grindstone**
Organized
Hardworking
Conscientious
Diligent

Trix & Psenka, Discourse and Society 14(2) 2003
Modified from Andrea Christopher, Alicia Carrasco and Shobha Stack
Clerkship Evaluations

Rojek et al JGIM 32(5) April 2019
Clerkship Evaluations

Non-URM and Men
Competency-Related Behaviors

- Advanced
- Clinical
- Compassionate
- Comprehensive
- Conscientious
- Efficient
- Knowledgeable
- Medical
- Scientific
- Thorough

URM and Women
Personal Attributes

- Cheerful
- Delightful
- Energetic
- Enthusiastic
- Lovely
- Mature
- Pleasant
- Reliable
- Respectful
- Wonderful

Rojek et al JGIM 32(5) April 2019
MSPEs

Non-URM =
Standout or Ability

Exceptional
Best
Superb
Intelligent
Brilliant

URM =
Competent*

Organized
Bright

* Secondary content analysis of use of “competent” found it to have a positive connotation less often when describing Black and Latino students than when describing Asian and white students.

Ross et al, PLOS ONE August 2017
The amplification cascade (Teherani, 2018 et al)

Figure 1 Fishbone diagram illustrating the causes, effects, and consequences of lower assessed performance in underrepresented in medicine (UIM) students compared with all students.
Disproportionate impact on UIM students impacts all students

**Faculty/resident rater factors**
- Variable teaching skills
- Variable assessment skills
- Insufficient performance observations
- Student mistreatment
- Personal burnout

**EFFECTS**
- Commission of microaggressions
- Unconscious bias
- Colorblind ideology

**CONSEQUENCES**
- Fewer honors grades and awards
- Fewer opportunities:
  - Career
  - Specialty
  - Residency program

Impacts all students
Disproportionate impact on UIM students
To be evaluated, learners need opportunities to perform

- Significant gender-based discrepancies in surgical case volumes among ENT residents
- Female residents reported 41 less cases than male residents per graduating year

Equitable opportunities for learning and evaluation are critical

Key Lessons from the Literature

• We all have implicit bias
• Bias can cause discrepancies in learning opportunities
• Bias is reflected in words that are used in assessment and evaluation
• Bias can affect how students are graded
• This effect is amplified and impacts careers long-term
Let’s think about words
Exercise

Using Narrative Language
What **word categories** do these words belong to?

**Word Categories**
- Standout
- Ability
- Grindstone
- Compassion
What word category does the word CARING belong to?

Standout
Ability
Grindstone
Compassion

Total Results: 0
What word category does the word INTELLIGENT belong to?

- Standout
- Ability
- Grindstone
- Compassion

Total Results: 0
What word category does the word DILIGENT belong to?

Standout
Ability
Grindstone
Compassion

Total Results: 0
What word category does the word STELLAR belong to?

Standout

Ability

Grindstone

Compassionate

Total Results: 0
Word categories commonly used

<table>
<thead>
<tr>
<th>Women/URM</th>
<th>Men/White</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grindstone</td>
<td>Standout words</td>
</tr>
<tr>
<td>Compassion</td>
<td>Ability</td>
</tr>
<tr>
<td>Personal attributes</td>
<td>Competency</td>
</tr>
</tbody>
</table>

Ross et al, PLOS ONE August 2017
Gender-bias calculator

This calculator is derived from the version made by Thomas Forth which was, in turn, inspired by this AWIS blog post on gender biases in recommendation letters. The blog post and the scientific paper it is based on also explain why this gender bias is important. Thanks to Dr. Karen James for the inspiration. Privacy note: no content you test here will leave your browser as all the calculation is done in this page.

Try an example!

Write or paste your recommendation letter here. Words that are more often associated with women will be added to the female list. Words that are more often associated with men will be added to the male list.

| Female-associated words | Male-associated words |

https://www.tomforth.co.uk/genderbias/
Use competency-based frameworks to minimize bias

- Encourage use of standardized language
- Anchors/milestones along the developmental trajectory
- Common understanding across institutions
Practice using a competency-based framework
Competency-based framework example 1:

AAMC’s 13 Core Entrustable Activities for Entering Residency²

1. Gather a history and perform a physical examination
2. Prioritize a differential diagnosis following a clinical encounter
3. Recommend and interpret common diagnostic and screening tests
4. Enter and discuss orders and prescriptions
5. Document a clinical encounter in the patient record
6. Provide an oral presentation of a clinical encounter
7. Form clinical questions and retrieve evidence to advance patient care
8. Give or receive a patient handover to transition care responsibility
9. Collaborate as a member of an interprofessional team
10. Recognize a patient requiring urgent or emergent care and initiate evaluation and management
11. Obtain informed consent for tests and/or procedures
12. Perform general procedures of a physician
13. Identify system failures and contribute to a culture of safety and improvement
Competency-based framework example 2:

Family Medicine Milestones

The Accreditation Council for Graduate Medical Education
## ACGME Family Medicine Milestone example

### Patient Care 3: Health Promotion and Wellness

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifies screening and prevention guidelines by various organizations</td>
<td>Reconciles competing prevention guidelines to develop a plan for an individual patient, and considers how these guidelines apply to the patient population</td>
<td>Identifies barriers and alternatives to preventive health tests, with the goal of shared decision making</td>
<td>Incorporates screening and prevention guidelines in patient care outside of designated wellness visits</td>
<td>Participates in guideline development or implementation across a system of care or community</td>
</tr>
<tr>
<td>Identifies opportunities to maintain and promote wellness in patients</td>
<td>Recommends management plans to maintain and promote health</td>
<td>Implements plans to maintain and promote health, including addressing barriers</td>
<td>Implements comprehensive plans to maintain and promote health, incorporating pertinent psychosocial factors and other determinants of health</td>
<td>Partners with the community to promote health</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comments:</th>
<th>Not Yet Completed Level 1</th>
<th>Not Yet Assessable</th>
</tr>
</thead>
</table>
Exercise 2: Pair-share

• For each comment on the slides that follow,

• Turn to your neighbor and discuss:
  • How do you categorize the language here?
  • How might you rewrite this comment?
  • Assume both students are demonstrating competencies at the highest levels and would be eligible for an honors grade at your institution.
(Student) was a dependable and strong member of the team. They responded well and improved with feedback given over the course of several directly observed encounters and demonstrated an exceptionally productive and positive attitude towards their education and professional development.
It was a delight to supervise (student) during this rotation. They were remarkably diligent, thorough, energetic, and dedicated to taking great care of our patients. I would gladly refer patients or family members to see them.
Exercise 2:

• Think of a learner that you recently worked with
  • If you really can’t think of one or don’t do clinical work, pair up!
• Write an objective, competency-based evaluation

QR codes to access competency-based frameworks

AAMC EPAs

ACGME FM Milestones
Maximizing Our Impact
Implementing Higher-Level Change to Reduce Bias

- Individual change is one way to reduce bias
- How can we scale up our impact?
- What is the institution's responsibility?
- Clerkship Directors & Program Directors: how can you use your power?
Impact Pyramid

- Individual Level
- Site Level
- School Level
- Structural Level

IMPACT

INDIVIDUAL EFFORT
Examples from Our Work

• Grading Committees
• LOR guidance with specific reference to reducing bias
• Diversifying faculty to reflect student experiences
Exercise 3: Small Group Work

Consider current teaching environment (clerkships, sub-Is, etc.) and brainstorm how to implement changes on a larger scale

• Start with thinking about your own role/institution
  • Faculty development, evaluation forms, etc.
• Form a group of 4 and share ideas
• Select a person to report 1 idea to larger group
Pause and Reflect

- One thing you'll take back home
- One thing you're still thinking about
Take homes

- TIP 1: Be aware that you have biases
- TIP 2: Know your learner as a unique individual
- TIP 3: Ensure equity with learning opportunities
- TIP 4: Slow down when writing an evaluation
- TIP 5: Use competency-based frameworks with specific examples
- TIP 6: Comment on knowledge and skills
- TIP 7: Check your words
- TIP 8: Consider system-level changes
Evaluation

Please complete an evaluation for this presentation.

Boatright et al. Racial Disparities in Medical Student Membership in the Alpha Omega Alpha Honor Society. JAMA Internal Medicine 2017.


Lee et al. 2007. “Making The Grade:” Non Cognitive Predictors of Medical Students’ Clinical Clerkship Grade. *Journal of the National Medical Association.*


Ross et al. Differences in words used to describe racial and gender groups in Medical Student Performance Evaluations. *PLOS one* 2017.


National Center for Women and Information Technology. Avoiding Unintended Gender Bias in Letters of Recommendation. [https://www.ncwit.org/sites/default/files/resources/avoidingunintendedgenderbiaslettersrecommendation.pdf](https://www.ncwit.org/sites/default/files/resources/avoidingunintendedgenderbiaslettersrecommendation.pdf)