2019-03-28: Bias in Examinations

The following links were shared during the chat:

- Relationships of demographic variables to USMLE physician licensing exam scores: a statistical analysis on five years of medical student data [https://doi.org/article/7f2b0f48b64c4e62827492203d9bf64](https://doi.org/article/7f2b0f48b64c4e62827492203d9bf64)
- Cultural Bias in University Entrance Examinations In the UAE [https://search-proquest-com.libproxy.lib.unc.edu/docview/1721972977?pq-origsite=summon](https://search-proquest-com.libproxy.lib.unc.edu/docview/1721972977?pq-origsite=summon)

---

**MedEd Chat @MedEdChat** 8 hours ago
Welcome to the #mededchat (US) I am your moderator for the next hour @alliance4clined #meded

**MedEd Chat @MedEdChat** 8 hours ago
The #mededchat topic & questions will be announced in a moment…for now, please introduce yourselves #meded

**Gary Beck Dallaghan @GLBDallaghan** 8 hours ago
@MedEdChat @Alliance4ClinEd Gary here in Orlando at the #SGEA2019 meeting #mededchat

**Dr. Christopher Tarolli @DrTarolliNeuro** 8 hours ago
@MedEdChat Chris here, neurologist in Rochester, NY #mededchat

**MedEd Chat @MedEdChat** 8 hours ago
@DrTarolliNeuro Welcome, Chris! #mededchat

**MedEd Chat @MedEdChat** 8 hours ago
If you are tuning in to the #mededchat, remember to use the #mededchat hashtag and try to answer with the Topic numbers (T1, T2, T3)

**MedEd Chat @MedEdChat** 8 hours ago
@AJKleinhex Happy to have you join, AJ! #mededchat
Brenda Roman  @BJBRoman 8 hours ago
#BrendaOnAStick checking in for #mededchat from #AAMcgea wearing my #garyonsocks  https://t.co/MP3YpBRYoc

MedEd Chat  @MedEdChat 8 hours ago
We will assume that all of your tweets during #mededchat are your own during this hour unless otherwise stated #meded

Jonathan Lim  @JonLimMD 8 hours ago
Jon, a BCM CMR in Houston, having a lot of FOMO from #SHM19 last week and #SGEA2019 this week! Excited to learn more about bias in assessment tonight with #mededchat

MedEd Chat  @MedEdChat 8 hours ago
TOPIC 1: What is meant by assessment bias? #mededchat #meded

Paul Haidet  @myheroistrane 8 hours ago
@BJBRoman Oh dear! #mededchat

R Klein MD MEHP  @RKleinMD 8 hours ago
@MedEdChat Looking forward to discussion on this topic. #mededchat

Gary Beck Dallaghan  @GLBDallaghan 8 hours ago
@MedEdChat T1 #mededchat I found some interesting discussions about this....from bias in test results to bias in test items. Which is more bothersome?

A.J. Kleinheksel  @AJKleinhex 8 hours ago
T1 Assessment bias is the cognitive and sociocultural biases we embed into our assessment instruments (usually unintentionally). #mededchat

Gary Beck Dallaghan  @GLBDallaghan 8 hours ago
T1 This was one article that considered bias in examinations  https://t.co/TlSEgrWwC4 #mededchat
Jonathan Lim @JonLimMD 8 hours ago
T1: My very basic understanding would be that assessment bias would be similar to any type of test we use in medicine that can have false positives or false negatives, issues with sensitivity and specificity #mededchat

R Klein MD MEHP @RKleinMD 8 hours ago
@MedEdChat Broad umbrella with this I think. Wondering how to prioritize types of bias in Assessment #mededchat

MedEd Chat @MedEdChat 8 hours ago
RT @GLBDallaghan: T1 This was one article that considered bias in examinations https://t.co/TlSEgrWwC4 #mededchat

Sateesh Arja, MD, MPH @ArjaSateesh 8 hours ago
@MedEdChat #mededchat T1 unfair grades. Assigning unfair grades or penalizing certain groups of students. I also consider being liberal or giving high grades which students don’t deserve can be considered as bias #meded

Jonathan Lim @JonLimMD 8 hours ago
T1 especially when so many of our assessments are psychosocial-- they have historical and cultural contexts-- usefulness for certain populations more than others #mededchat

Gary Beck Dallaghan @GLBDallaghan 8 hours ago
@AJKleinhex T1 So you're looking at it more from an item level bias? How do you think that impacts the learners? #mededchat

Dr. Christopher Tarolli @DrTarolliNeuro 8 hours ago
@GLBDallaghan @MedEdChat #mededchat think bias in test items more insidious and potentially more dangerous because more challenging to identify

Alliance4ClinEd @Alliance4ClinEd 8 hours ago
T1 Now that is an interesting take on bias with exams. Would you say that OSCEs can have that kind of bias due to elevated grades? #mededchat

Gary Beck Dallaghan @GLBDallaghan 8 hours ago
@DrTarolliNeuro @MedEdChat Tell me more. Why do you think item level bias is more dangerous? #mededchat
A.J. Kleinheksel @AJKleinhex 8 hours ago

@GLBDallaghan T1 Well, in that instruments are made up of items. But bias can be found in both individual assessments as well as systems (and constructs) of assessment. Bias affects learners by unfairly putting them at a disadvantage due to their identity &/or socioeconomic status. #MedEdChat

Dr. Christopher Tarolli @DrTarolliNeuro 8 hours ago

@GLBDallaghan @MedEdChat I feel like bias in test results may be more obvious, so therefore easier to identify and correct. Bias in items much less likely to be identified by students or educators - so potentially more dangerous. #mededchat

MedEd Chat @MedEdChat 8 hours ago
RT @AJKleinhex: @GLBDallaghan T1 Well, in that instruments are made up of items. But bias can be found in both individual assessments as we…

Gary Beck Dallaghan @GLBDallaghan 8 hours ago
@AJKleinhex You’re hitting at it from a macro level down to the individual. Nicely stated. #mededchat

MedEdBot @MedEdBot 8 hours ago
RT @MedEdChat: TOPIC 1: What is meant by assessment bias? #mededchat #meded

Dr. Christopher Tarolli @DrTarolliNeuro 8 hours ago
RT @AJKleinhex: @GLBDallaghan T1 Well, in that instruments are made up of items. But bias can be found in both individual assessments as we…

Larry Hurtubise @hur2buzy 8 hours ago
#mededchat checking in from #aamcCGEA with @2LindaMLove @carens8892 #garyonsocks https://t.co/wljzMPmGAA

A.J. Kleinheksel @AJKleinhex 8 hours ago
@GLBDallaghan ALL THE LEVELS #mededchat https://t.co/IyU6uKuvuy
Sateesh Arja, MD, MPH @ArjaSateesh 8 hours ago
@AJKleinhex #mededchat T1 every assessment has hidden bias as we are assessing the proxy of learning and we are not actually measuring what is learned. Some assessments require written and verbal communications. It can be a bias if students are not taking these exams in their native language.

MedEd Chat @MedEdChat 8 hours ago
RT @AJKleinhex: @GLBDallaghan ALL THE LEVELS #mededchat https://t.co/IyU6uKuvuy

Paul Haidet @myheroistrane 8 hours ago
If humans are involved, there’s potential for bias, either in the construction of the instruments themselves (think SAT) or in the application of instruments (think differential application of rubrics) #mededchat

MedEd Chat @MedEdChat 8 hours ago
RT @myheroistrane: If humans are involved, there’s potential for bias, either in the construction of the instruments themselves (think SAT)…

Paul Haidet @myheroistrane 8 hours ago
@hur2buzy @2LindaMLove @carens8892 @GLBDallaghan Where can I get a pair of #garyonsocks?? #mededchat

ADMSEP @admsep 8 hours ago
@ArjaSateesh @AJKleinhex That’s an excellent point. The psychological toll that can take with oral exams also interferes with performance. #mededchat

Miranda Huffman, MD, MEd @mirandahuffman 8 hours ago
@MedEdChat It’s easy to give better grades to students you like or can relate to. #mededchat

Gary Beck Dallaghan @GLBDallaghan 8 hours ago
@mirandahuffman @MedEdChat In those situations, how do you check yourself to not be biased? #mededchat
Sateesh Arja, MD, MPH @ArjaSateesh 8 hours ago

@Alliance4ClinEd #mededchat T1 to some extent yes. Even though OSCE, we use checklists or rubrics there is still possibility for bias. If we are using global ratings, there is more possibility for subjective bias. But grades by checklists and globally ratings are generally comparable #meded

MedEd Chat @MedEdChat 8 hours ago

RT @ArjaSateesh: @Alliance4ClinEd #mededchat T1 to some extent yes. Even though OSCE, we use checklists or rubrics there is still possibil...

MedEd Chat @MedEdChat 8 hours ago

TOPIC 2: If bias exists, what does that say about the validity of the assessment? #mededchat #meded

Jonathan Lim @JonLimMD 8 hours ago

T1 in UME/GME, I think end of rotation evaluations that we currently use can be challenging. The 1-10 can be interpreted differently, but in particular, I worry about the narrative assessment that the teacher does of the student as being vulnerable to unconscious bias #mededchat

Paul Haidet @myheroistrane 8 hours ago

@mirendahuffman @MedEdChat In the communications literature, Lisa Cooper found that race concordant pairs spent longer with each other, often spending the time on social chit-chat. Relationship building is better when there’s a connection - probably in assessment, too. #mededchat #https://t.co/7aUdlc2RV

A.J. Kleinheksel @AJKleinhex 8 hours ago

@JonLimMD @GLBDallaghan Certainly! Item-level assessment bias may be implicit (e.g., an example in a question that only upper SES would get, such as referencing certain financial products), or explicit (e.g., racial stereotypes in a case vignette). #MedEdChat

R Klein MD MEHP @RKleinMD 8 hours ago

@MedEdChat Assessment validity should be what we strive for. But its not a given. #mededchat
In the communications literature, Lisa Cooper found that race concordant pairs spent longer w…

T2 a consistent bias probably is an indictment of the construct validity of the assessment tool-- it's measuring some other factor that's impacting the results #mededchat

RT @JonLimMD: T1 in UME/GME, I think end of rotation evaluations that we currently use can be challenging. The 1-10 can be interpreted diff…

Agree! I think we often are happy to accept a Faldo sense of security in assuming assessments are free of bias.#mededchat

T2 validity is more of accuracy of the assessments rather than results in my opinion. Validity is more dependent on content of the assessments, blueprinting, learning objectives and outcomes. Bias distorts more of reliability which is reproducibility of assessments.

System-level bias could be the inclusion of problematic constructs (e.g., “professional attire” that penalizes natural hair styles), or course grading that is affected by personal instructor bias, or institutional frameworks that assume access/resources. #MedEdChat

T2: we can only know how our instruments (tests, OSCEs, clinical forms, etc) perform for different groups if we look. Examining differences along ethnicity, gender, disability, etc should be a routine part of analysis, and should be investigated when differences found. #mededchat
RT @myheroistrane: T2: we can only know how our instruments (tests, OSCEs, clinical forms, etc) perform for different groups if we look. Ex…

Gary Beck Dallaghan @GLBDallaghan 8 hours ago
T2 Even though there is evidence of poor performance by certain groups on the MCAT, the authors of this article felt there really was no bias in the exam. Do Racial and Ethnic Group Differences in Performance on the MCAT Exam Reflect Test Bias? https://t.co/My3Q77Utrg #mededchat

A.J. Kleinheksel @AJKleinhex 8 hours ago
@JonLimMD Agreed! If there is demonstrable bias, the instrument cannot be valid. #MedEdChat

Gary Beck Dallaghan @GLBDallaghan 8 hours ago
@ArjaSateesh @MedEdChat I don’t know. There are aspects of validity evidence that are impacted by bias in exams. Consequences validity as well as response process are pieces of validity evidence that are compromised by bias. #mededchat

A.J. Kleinheksel @AJKleinhex 8 hours ago
@ArjaSateesh @MedEdChat I don’t know, I think an instrument (and/or person) can be consistently biased. Construct validity is where I think that bias is typically found. The content included maybe valid, but the construct is off-center. #MedEdChat

Nav Sidhu @NavSidhu08 8 hours ago
RT @myheroistrane: T2: we can only know how our instruments (tests, OSCEs, clinical forms, etc) perform for different groups if we look. Ex…

Nav Sidhu @NavSidhu08 8 hours ago
RT @JonLimMD: T1 in UME/GME, I think end of rotation evaluations that we currently use can be challenging. The 1-10 can be interpreted diff…

Jonathan Lim @JonLimMD 8 hours ago
T2 what if all our assessments are biased and therefore not valid. help! #mededchat
https://t.co/y8cywJVTTH
**MedEd Chat** @MedEdChat 8 hours ago
T2 How would you address threats to validity with biased items? #mededchat

**Paul Haidet** @myheroistrane 8 hours ago
It’s hard to stay focused when the game is tied with 0.1 sec left! #mededchat
https://t.co/k0WNiTceji

**A.J. Kleinheksel** @AJKleinhex 8 hours ago
@GLBDallaghan I’m paywalled in the hotel, but I suspect this would be a case of systemic bias, rather than item-level. Is that the case? #MedEdChat

**Gary Beck Dallaghan** @GLBDallaghan 8 hours ago
@AJKleinhex T2 #mededchat I think it is. Reading some of the letters rebutting their findings gave me hope that people didn't buy their conclusions.

**MedEd Chat** @MedEdChat 8 hours ago
TOPIC 3: What are strategies to ensure assessments are not biased? #mededchat #meded

**Matt Lineberry** @TestTransformed 8 hours ago
@ArjaSateesh @MedEdChat #mededchat T2 We generally say “bias” to mean a consistent, “directional” error in scores (a validity issue); reliability is all about - and only about - random (non-directional) error

**MedEd Chat** @MedEdChat 8 hours ago
RT @TestTransformed: @ArjaSateesh @MedEdChat #mededchat T2 We generally say “bias” to mean a consistent, “directional” error in scores (a v…

**A.J. Kleinheksel** @AJKleinhex 8 hours ago
T2 As with all instrument design, there are methods to ensure rigor and fairness. But those processes have to be inclusive if we want our assessments to be inclusive: diverse panels for expert review, pilot testing across all appropriate demographics, etc.
#mededchat
Jonathan Lim @JonLimMD 8 hours ago
T3 probably careful #DiversityandInclusion in populations used for assessment validity testing such as think-alouds or more broader statistical validation. Probably also important with iterative faculty revision of an assessment tool that they are a diverse set #mededchat

MedEd Chat @MedEdChat 8 hours ago
RT @JonLimMD: T3 probably careful #DiversityandInclusion in populations used for assessment validity testing such as think-alouds or more b…

MedEd Chat @MedEdChat 8 hours ago
RT @AJKleinhex: T2 As with all instrument design, there are methods to ensure rigor and fairness. But those processes have to be inclusive…

Paul Haidet @myheroistrane 8 hours ago
RT @TestTransformed: @ArjaSateesh @MedEdChat #mededchat T2 We generally say “bias” to mean a consistent, “directional” error in scores (a v…

Jonathan Lim @JonLimMD 7 hours ago
T3 I feel the bias revealed especially when I've made a survey and then give it to someone even a few years my junior and they have no idea what I'm talking about or in the reverse to a senior faculty member as well. Or it's just me writing bad questions... or both #mededchat

Gary Beck Dallaghan @GLBDallaghan 7 hours ago
@JonLimMD As you and @AJKleinhex noted, pilot testing and using reaction panels of diverse populations is important to address bias #mededchat

Paul Haidet @myheroistrane 7 hours ago
T3: increase diversity among faculty; examine concordance between assessor and assessee and look for concordance effects in clinical assessments. #mededchat

MedEd Chat @MedEdChat 7 hours ago
RT @myheroistrane: T3: increase diversity among faculty; examine concordance between assessor and assessee and look for concordance effects…
Matt Lineberry @TestTransformed 7 hours ago
@ArjaSateesh @MedEdChat #mededchat T2 One trick with bias: you can’t sample your way out of it. A very smart #meded scholar was once pushing “wisdom of the crowds”: “just ask 40 people, ‘would you let this doc treat your mom?’” But that does almost nothing to mitigate common biases.

Miranda Huffman, MD, MEd @mirandahuffman 7 hours ago
@MedEdChat Part of the way to increase validity is #facultydevelopment on the instrument. Are our #meded faculty open to acknowledging their biases? #mededchat

MedEd Chat @MedEdChat 7 hours ago
RT @TestTransformed: @ArjaSateesh @MedEdChat #mededchat T2 One trick with bias: you can’t sample your way out of it. A very smart #mededsch...

Paul Haidet @myheroistrane 7 hours ago
@mirandahuffman @MedEdChat I have my reservations... #mededchat

Miranda Huffman, MD, MEd @mirandahuffman 7 hours ago
@MedEdChat T3: best way to avoid bias is having multiple evaluators and multiple assessments. Avoiding any high-stakes assessments. Portfolios of work. #mededchat. #step1passfail

A.J. Kleinheksel @AJKleinhex 7 hours ago
T3 The elimination of assessment bias is bidirectional from the instrument to the system. Eliminating bias in both requires greater transparency, increased rigor, more frequent evaluation, and many more open and honest conversations. #mededchat

Paul Haidet @myheroistrane 7 hours ago
RT @AJKleinhex: T3 The elimination of assessment bias is bidirectional from the instrument to the system. Eliminating bias in both requires...

MedEd Chat @MedEdChat 7 hours ago
RT @AJKleinhex: T3 The elimination of assessment bias is bidirectional from the instrument to the system. Eliminating bias in both requires...
Miranda Huffman, MD, MEd @mirandahuffman 7 hours ago
@myheroistrane @MedEdChat Me too. It’s hard to admit your own biases. #MedEdChat

Gary Beck Dallaghan @GLBDallaghan 7 hours ago
@mirandahuffman @myheroistrane @MedEdChat T3 If we set the example of being transparent that gets the ball rolling. Most people who know me know I will put that out there just to ensure there is understanding of where I’m coming from. #mededchat

Paul Haidet @myheroistrane 7 hours ago
@mirandahuffman @MedEdChat Agree. First step in mitigating bias: faculty need to “own” their personal biases. #mededchat

MedEd Chat @MedEdChat 7 hours ago
We have about 5 more minutes left in our discussion. Please feel free to give some final thoughts #meded #mededchat

Gary Beck Dallaghan @GLBDallaghan 7 hours ago
@myheroistrane @mirandahuffman @MedEdChat There is also a need to respect the others who the bias may be intentionally or unintentionally leveled at. #mededchat

Jonathan Lim @JonLimMD 7 hours ago
Just want to share a moment of gratitude-- nothing like a fun, engaging #mededchat with experts that leads to a lot of learning and thought-provocation! Feeling uplifted and inspired #mededchat #meded

Gary Beck Dallaghan @GLBDallaghan 7 hours ago
Final thoughts: This issue of bias in testing was new for me. I appreciate @AJKleinhex for suggesting it. I feel like I’ve learned a good deal over the past hour. #mededchat

Matt Lineberry @TestTransformed 7 hours ago
@ArjaSateesh @MedEdChat T2 #mededchat It is easy to forget that admittedly limited test methods (MCQs and the like) were once a pretty remarkable social justice innovation: moving away from the subjectivity in oral exams/etc., when examiners may not have believed some types of people “should” be docs
Join us again next week at 9 pm Thursday. Don't forget to suggest topics by DM or email #meded #mededchat

All kinds of diversity are needed. Our faculty should mirror the population. #MedEdChat

T3 what if the bias come from a SP rather than a faculty member in case of OSCE? I think constant analysis of results and data is required to find out if there is any bias. Evaluation of SP, station itself, and faculty required. Training of faculty and SP is required.

Final thought: to mitigate bias: 1) “own” yours 2) develop situational awareness of bias activating 3) call a time out when it does 4) ask yourself how you would act if the other person were your ethnicity/gender/ability/etc 5) repeat liberally #mededchat

That’s a wrap...I will post the #mededchat transcript tomorrow morning on [https://t.co/mJivoKroXx](https://t.co/mJivoKroXx) on the Resources page. Thanks everyone for participating! #meded

RT @myheroistrane: Final thought: to mitigate bias: 1) “own” yours 2) develop situational awareness of bias activating 3) call a time out…

RT @myheroistrane: Final thought: to mitigate bias: 1) “own” yours 2) develop situational awareness of bias activating 3) call a time out…

agree. I too learned many new things.
RT @AJKleinhex: T2 As with all instrument design, there are methods to ensure rigor and fairness. But those processes have to be inclusive…

RT @mirandahuffman: @MedEdChat T3: best way to avoid bias is having multiple evaluators and multiple assessments. Avoiding any high-stakes…

RT @AJKleinhex: T3 The elimination of assessment bias is bidirectional from the instrument to the system. Eliminating bias in both requires…

And willingness to develop and empower junior faculty, esp those who don’t look like traditional academic medicine faculty. #MedEdChat

RT @myheroistrane: T3: increase diversity among faculty; examine concordance between assessor and assessee and look for concordance effects…

Great point. I also think about communication - are our assessment tools assessing how a white male physician would best communicate with a white, educated, sober, heterosexual patient? Is this why we’re having so much trouble impacting #sdh? #MedEdChat

Or be vulnerable and ask someone else who doesn’t share your lived experiences to call you out. #MedEdChat

How much is the purchase price? Can I buy in bulk? #GaryOnSocks#MedEdChat

Geographic region, age, SES, language spoken at home. So much diversity in the world! #MedEdChat
RT @mirandahuffman: @TestTransformed @MedEdChat @ArjaSateesh Great point. I also think about communication - are our assessment tools asses…