(PLEASE NOTE: Transcribed automatically by Vimeo; mistakes are possible/likely. Our apologies.)

TRANSCRIPT - GR 01 31 25 *"**Fair or Flawed? A Critical Look at (in) Equity in Assessment**"* guest speaker Benjamin Kinnear MD MEd, University of Cincinnati

Internal Medicine Grand Rounds

- We're delighted. We're delighted here today to be hosting Dr. Ben Kinnear for a virtual grand rounds, speaking about fair or flawed a critical look at inequity in assessment.
- I'll take us through our Cma disclosures and we're faculty claiming credit. The activity code for today's session is here as well as the methods for claiming credit, and the email will land in about 30 min and next pass over to Dr. Shaina Hassan to introduce Dr. Munich.
- So good afternoon. Everyone. It is my pleasure to introduce our virtual grand round speaker for today a little bit of background about him. Dr. Piner obtained his medical training at the University of Missouri, and then comes on to a Med Residency as well as Chief Residency at the University of Cincinnati Medical Center and Cincinnati Children's Hospital, where he's now an associate professor of internal medicine and pediatrics.
- He's the current program director for the Internal Medicine Pediatrics Residency Program helping to lead Quality improvement curriculum. The Resident medical education based assessment initiatives. In addition to serving as a program director for the Imsr Medical Education website.
- Dr. Janeer also obtained his master of Medical education from the University of Cincinnati in 2018, and completed a 1-year research fellowship with the Education Research Scholarship program at Cincinnati.
- At that time. He was also selected for the Macy faculty scholars program during which he piloted competency-based time variable training in the Uc. Internal medicine. He's currently a Phd. Student at Masters University, where he is studying liberty, argumentation, and argumentation.
- On a personal note. Dr. Kinir spends most of his free time with his wife and 2 daughters hiking, playing golf games, and traveling whenever possible. He's a St. Louis cardinals, fanatic, and strongly believes that mint flavored ice cream is an evolving nation. So we are excited to have us for grandmother. So please join me in welcoming Dr. Kinir, and hopefully he can tell us a little bit more about his dislike for mint ice cream.

Benjamin Kinnear (he/him)

While they're getting that set up. I will say I understand that right now, particularly the topic of equity can be kind of a hot button issue. It's been a pretty fraught discussion in political circles and things. But me bashing mint ice cream might be the most controversial thing I'm going to say today.

- And happy to go toe, to toe with anybody in the chat or in questions afterwards. If you want to try to defend the honor of mint ice cream, I think it's terrible but that's

not what we're here to talk about. We're here to talk about equity and assessment, and I'm really grateful for the opportunity to come chat with you all today

- I'm really thankful for the invitation and for everyone taking their time. I think this is a really important topic, and what I'd like to do is walk us through. What do we even mean by equity? What's the evidence of the problems in equity and assessment? And then what is there that can maybe help us move forward with the discussion.
- Here we go. Here are my learning objectives. I think these were sent out ahead of time. Probably, so I'm not going to read them to you, but I do want to talk about some disclosures so I don't have any relevant financial disclosures. But I have a couple of other disclosures. One is that I am an absolute assessment. Nerd. Most of my non-clinical time is spent thinking about assessment, and how can we make assessment more valuable for learners and for programs and frankly for our patients?
- But when it comes to equity, I'm still very much a learner. I am not an equity scholar. I am not a critical theorist but this topic is becoming increasingly important in the world of assessment. It's always been important. But it's really the growing course of people saying we need to do better is getting louder. And so I'm really grateful to have the space to talk about it with you all, and my last disclosure is, I just want to recognize my positionality in this whole discussion. So I am a cisgender, heterosexual, able-bodied white man and my identity certainly doesn't prevent me from experiencing or seeing inequity in the world, but it comes with its own set of privileges and blind spots, and I just wanted to acknowledge the fact that equity is not a like totally neutral topic. It's a it's a very fraught topic, and a very charged topic, and many of you may have certain experiences with inequity or bias or discrimination that I may not have had to deal with in my life, and so just wanted to throw that out there in terms of where I'm coming to this discussion from.
- So how does an assessment nerd like me, really. Start to gravitate toward the issues of equity that we have an assessment. It's not just me. A lot of assessment. Folks are moving in this direction to try to figure out, how can we do better? And for my personal story, it really started about 15 years ago, when our internal medicine program director. His name is Eric Warm said. We need to just kind of reset reinvent, rethink our entire assessment system. This is right around the time that competency based. Education is coming into the Zeitgeist of Med. Ed. And he said, Let's blow it up. Let's try to do much better.
- And this sent us on this long decade and a half journey that we're still on to improve assessment in our system, really focusing on defensibility and validity which are terms. You'll hear a lot in assessment. And so, just as some examples of things that we did, we completely restructured the way rounds happen in our hospital to get everything to the bedside, so that our clinical supervisors are right there with our learners watching them interact with patients much more frequently to actually do direct observation for assessment.
- We realigned all of our assessment instruments to use an entrustment supervision scale, and provided a lot more faculty development around how to use those things.
- We recognize that there was a lot of context, specific factors that influence ratings. And so we developed a set of learning analytics that help us take any piece of data that we get and put it in context in terms of who was the rater? What time of year was it? What rotation are they? On? What questions were asked, so that we can try to find signal amongst the noise and and try to really find ways to to find meaning in

the data, and we put these on dashboards. And now have it. All these analytics are integrated throughout our system.

- We started using clinical performance measures. So our internal medicine residents have a unique learning experience where they are embedded in an ambulatory long block for an entire year, and they are in clinic multiple times per week every week for an entire year with the same panel of patients, and we follow some of their process and outcome measures to see hours, their doctoring, impacting the care of those patients.
- We have another 3 60 assessment system where we excuse me, we gather data from patients, from peers, from staff, from faculty and from the learners themselves.
- We have a homegrown testing program that has been shown to really improve the scores of our learners. But even more importantly than that we can track. How are they engaging with those learning resources? Kind of taking a master adaptive learner approach of how are you using the the resources we give you to deliberately close gaps in your knowledge and skills.
- We follow things like, Are you closing your charts on time? Are you doing your refills on time? Are you getting all your modules and things done? We call those our citizenship measures? And then, lastly, we have a lot of narrative that we collect and have spent a lot of time building theories around. What do the different types of narrative mean to try to draw conclusions both to give feedback and to decide if and when learners are ready.
- So a lot of time, everything that I just summed up in one slide is over 10 years of work, lots of research put into it. Lots of thought, lots of time spent in building validity arguments around those data. But we didn't just want to stop with collecting data. We also wanted to try to make good decisions with the data.
- And so we dove way into the literature to find out what are the best practices around, how things like competency committees or entrustment committees, or grading committees, or whatever kind of group you have making decisions. How can we try to use the best theories and approaches possible to make defensible decisions? We put all this thought into it, and what we do is still very, very flawed. I know y'all do amazing assessment work there. You probably have a similar feeling. You pour all your time into it.
- It's still flawed, but it's better than where you started, which is, which is the best you can hope for.
- And then this thing happened where, a few years ago the Macy Foundation reached out. If you're not familiar with the Macy Foundation, it's a it's a really cool organization that funds a lot of medical education, research and innovation work, but it also has a conference once or twice a year, where it brings national experts together and tries to kind of find a way forward around a difficult topic, and they reached out and they said, Hey we're going to have a conference on fairness and equity in assessment. So back in 2021, I believe, and they said, Hey, we hear you have a really great assessment system there that you're doing some cool stuff. We'd like to invite a couple of you to come, we said, Yay, that's great. And they said, and in preparation we'd like for you to write a case, study paper about your assessment system, and we were like, Woo, we're so good we're doing great. And then they said, we said, Sure, we'll write that paper.
- And they said, awesome. The case study we want you to write is how you built an equitable assessment system.
- And then we were like, Oh, hmm! It's about the equity of our assessment system.

- I don't know that we can write that paper. We had spent so much time thinking about validity and defensibility, and certainly equity is something we valued, but as we designed our system and our approaches, I have to be honest. We probably well, we certainly were not nearly as intentional about equity as as we thought we were, and if we were to write a paper talking about how equitable our system was that was going to be really disingenuous. So we were kind of at this crossroads and didn't know what to do. And I want to put a PIN in that moment, because I'm going to come back to it. But the short story is, we said, we can't write that paper, not in good faith but we'll come back to it. Let's shift to the topic of equity, and how it ties into assessment. And I think one of the most foundational things I hope you take away from this talk is how is it like, what do you think equity is? And how do you think about it?
- It's a really complex construct. There's not a single definition that is universally accepted. And that's okay.
- Many people start off with a picture something like this. And I think it's because we often conflate equality and equity. And these kind of graphics are really helpful in drawing the distinction. So equality is when everybody has the same thing, and equity is when everybody gets what they need to succeed or to thrive, and sometimes you'll see various versions of this, or sometimes they'll expand it out to include justice or other concepts that are kind of helpful to try to see where there's comparisons and contrasts.
- But in terms of actually integrating this idea into assessment. You have to get even a little more nuanced about. What are we even talking about with equity and assessment? And there's lots of definitions out there. This definition comes from the report that stemmed from that same conference I was just talking about. So after they held the conference they write a report, and the Macy Foundation says they think equity is recognizing that we do not all start from the same place because power is unevenly distributed.
- I think it's a pretty pretty good definition.
- There's another paper that was led by Dr. Catherine Lucy, a really brilliant Med. Ed. Researcher, and they wrote that equity is the state of being fair and achieving social justice and that's also pretty good definition.
- In the same paper later, Dr. Lucy says, equity is when all learners have fair and impartial opportunities. Now to me, all of these, all of these definitions like have some truth to them, and they resonate with me to some degree. But they're all different. They're all describing somewhat different things. And I think it's important to recognize why that is, it's not because there's a problem with equity, or that there is a it's not. It's not an issue. It is a normal thing that happens anytime. You have a complex concept or construct that travels and migrates across different fields and domains and communities.
- And so equity, for example, or educational equity, or equity of assessment that is not something that originated in medical education or health professions, education. It's obviously a concept that's been around for a long time and is used in other fields. I mean, equity is a concept that is pretty universal and can be used almost anywhere and the point is that every time something like equity migrates across ideologies or fields, it gets recontextualized, and to a certain degree changes a little bit, and medical education itself is not monolithic. So it makes sense that in different spaces of Med. Ed, people might think about it differently, or use slightly different definitions. The one issue is what Dr. Edgar brings up in his paper that when it is

recontextualized it can lead to equity, being used uncritically to refer to a range of related and sometimes contradictory concepts. It's not an issue that we have different ways of thinking about it. But when we're not critical about it, when we're not clear, it can really be difficult to have generative conversations about what we mean, especially when you're trying to tackle like practical problems, or you're trying to make real world improvements. It's hard to set those goals and move forward. If we're not even talking about the same thing sometimes, and one of the best papers that I've come across. One of the most helpful, clarifying resources is this paper that was led by Hannah Anderson, who's an education researcher at Children's Hospital, Philadelphia. She's absolutely brilliant. In fact, everybody on this paper is brilliant. Great powerhouse author team and Hannah uses this, I think really helpful analogy. When thinking about equity, she uses the analogy of a mountain, and she says, You know, when you look at a mountain.

- There's not just one way of looking at it. If you were standing on the north side of the mountain, it'll look one way. You're standing on the east side of the mountain, it might look slightly different. If you're standing on different elevations, it'll look different and different perspectives. So if you're looking at the Devil's Tower in Wyoming. It looks like this when you're standing a certain distance away, and it's this bright, beautiful orange, and the striations going down. It are nice and straight, and the trees below it are really green. It looks like it's surrounded by like a lush forest but if you get up closer, you lose some of that forest. It looks a little more rocky at the base, actually no longer looks orange. It looks kind of like a faded yellow, or maybe whitish. The cragginess at the top is more pronounced than you thought. Just it looks totally different if you change your perspective, and if you look at it from this perspective, it's even more different. You're looking down at it just looks like a circle with a big shadow behind it.
- These are all Devil's Tower in Wyoming, but depending on your perspective or your orientation, you might have a very different description or understanding, or different things that you emphasize about it. And and Hannah really beautifully writes that equity is the same way and that people have different orientations when approaching equity, and each of those orientations brings a different set of assumptions or priorities or understandings with it, and none of them is right or wrong. But what's helpful is to try to clarify. What do we even mean when we're saying equity? Do we have a certain orientation that we bring to it? And can we be transparent about that?
- She writes about 3 different orientations, the fairness, orientation, inclusion, and then justice. And what I want to do is just very briefly go through each of these. I'd like for you to think about. Is there one that resonates with you, or do they all resonate with you? And is there a way that you can be clearer or be clear in general when you're having conversations about equity, or you're trying to do equity trying to do or to promote equity about what orientation you're coming from and how do you solicit that from other people?
- So the fairness, orientation is really about equal opportunity and being impartial. The idea is that everybody has an opportunity to be assessed and to have learning experiences. And we're going to remove bias. We're going to remove anything that feels unfair and be totally transparent about our process. This is what you would call the fairness orientation, and I think for a long time in medical education. This is what most people thought of as as equity, and may still be the predominant view depending on where you are.

- The inclusion. Orientation is an orientation that is a really strong narrative, particularly in literature and discussions around learners with disabilities. And it's about, how do we make our systems adaptable and supportive of all learners, even if they have different needs. So it's not about everyone gets an equal opportunity. Necessarily you recognize that people might need different opportunities, and you might need individualization and flexibility of the system to meet people with their needs where they are, and it really leans heavily into co-design, because the people who might need individualized education and assessment are the ones that need to help us co-design those systems. So again, that's the inclusion, orientation and the last one is the justice orientation.
- What the justice orientation says is that trying to get toward impartiality is probably a bit of a fool's errand, because nothing is impartial and it the justice. Orientation says that everything in our systems are tied to a broader historical and cultural and social and political systems, and you can't divorce the work that we do from those things you have to acknowledge them. You have to engage with them, and only by like going that extra layer, deeper, and looking at those external drivers that impact our system? Can you really make change from an equity standpoint? And so sometimes you have to tear down parts of, or all of the current systems and build it back with the most minoritized and marginalized people in our communities leading the way. This is the justice orientation.
- And you might be saying these all sound, really good. I don't think anyone is the right one, and you would be correct. There's not. This is not a competition. This is not an either, or they are just different again, orientations that people tend to bring to conversations. But we don't always put forward the orientation we're bringing to those conversations, and it can lead to tension. And it can actually be an obstacle to trying to move forward equity work in the assessment space. And this is another figure from Hannah's paper that I think is really helpful to show that if each of these circles represents one of those orientations you'll see that they overlap. They're not separate circles. They have a lot of overlap and the thing that I also want to emphasize is that in the middle of this this kind of gray Venn diagram.
- Hannah writes that all of them conceptualize equity as having both a process component and an outcome component.
- And so, in addition to having you reflect on your orientation that you bring to equity, or that you have other people like the orientations you see in the work that goes on around you. I also want to call attention to the fact that we need to consider both processes and outcomes of assessment as we try to improve the equity of what we're doing most of the time, I think people gravitate toward outcomes. But the processes are just as important. So thinking about in terms of equity, whether you take a fairness and inclusion, or a justice orientation, what are the tools that you use? What are the assessment instruments that you use in your system? What are the processes you use to gather data?
- What's your learning environment like? Do people have identity? Safety? Is? Are there microaggressions or macroaggressions? Is there implicit bias, equal opportunities or equitable opportunities to to engage in learning?
- And then, once you collect all that data, how are you sharing it with people, both in terms of making summative decisions and sharing those decisions with learners? But also, how are you using your assessment data to help people get better and to coach them and help them grow. Those are all process measures. And again, these are these are really important to consider when you're trying to figure out how

equitable is my assessment system. And then, of course, there's the outcome measures that many of us think of reflexively like. Are there differential achievement between identities and groups of learners in terms of their grades or rankings, or scores the awards that they might get, or the opportunities that people have in the sponsorship opportunities.

- Both of these are really important. And you know, I this might just be my Qi brain. But the reason I like holding both of these together is because a lot of times in quality, improvement or systems improvement, we spend a lot of time building a theory of change to try to understand. How do the processes that we have linked to the outcomes that we care about? And if you have a good theory of change. You can then intervene on those processes to try to change the downstream outcomes, and I think we can approach equity and assessment in a similar way, building theories of understanding, of how the processes, impact the outcomes and then targeting our interventions. And I'll show you in just a little bit an example of a program that did just that which I think was really really cool.
- So in terms of processes and outcomes of assessment, how are we doing in terms of equity. This is like the one, maybe not the one but one of the womp womp slides that I want to share, which is not so good.
- You probably saw this coming based on the title of my talk, but, as far as I can tell there is not a single general approach to assessment that has that has been shown to not have bias or evidence of inequity, either at the process level or the outcome level.
- Whether it's observation based assessment where we observe our learners on rounds and give them ratings, or write comments about them, or standardized exams like osces or multiple choice tests, or even, you know just about anything actually has been shown to have issues with inequitable outcomes or processes. And this has been shown across Ume and Gme, I'm not going to do a tour to evidence, because I'm giving you the summation of it right now, and it's kind of a really sad and growing body of literature that is getting harder and harder to not feel a strong urgency to fix, and I'll make an argument for why, that is in just a second.
- But I also want to point out this is not just like when I say assessment. I don't just mean like the scores people get and the narrative. We write about them for assessment forms and things. It's even for the downstream consequential things, too, like letters of recommendation which we may not think of classically as assessment. But it's a form of assessment that has massive ramifications for a lot of our learners. There's a ton of literature showing that the language that we use about our learners varies widely based on people's identities and backgrounds. Just as a simple example is, multiple studies have shown that if you are a gender minority, or a woman, or if you come from a minoritized race or ethnicity background, that people are more likely to use communal terms in your letter of recommendation when writing about you, as opposed to agentic terms, which are used more frequently with men or with white applicants.
- Now, some of you have probably heard this before, but if you haven't, you might be saying those communal terms look really great like, I want people to come to my program who are warm and empathic and helpful, and team players. And that is true. Nobody is saying that agentic is better than communal, or vice versa but it has been shown repeatedly that when you have agentic language in something like a letter of recommendation, that those people tend to get better outcomes in terms of seeking the jobs and positions they want. So it's an advantage for applicants to

have more agentic language in their letters of recommendation. And we differentially use this language based on people's gender identities and races and ethnicities. And there's other things that we do, too, like hedging and faint praise, and all sorts of other linguistic tricks that we do, that we apply inequitably in our letters of recommendation. So it's not just an assessment issue like during training. It's also in this transition period where so much pressure hinges on things, we're not doing a very good job.

- And hopefully, if you if you hear this, you're thinking this is a moral issue. We should not be treating people in inequitable manners. But there's also real world, tangible consequences for our learners and for our patients.
- One of the things that Dr. Ariana Tehrani, who's at Ucsf, says, is that there's an amplification cascade that happens over time, even though each one of these moments of inequity might seem small, like slight differential use of narrative on an assessment form, or a slight tweak of a quantitative score based on someone's identity or background. It seems small, but it amplifies over time and leads to these big differences downstream, such that inequitable assessment, even in small moments, can over time, lead to differences in how many honors you get, or your class rank, or all these other ways that we use to stratify people.
- It also impacts whether you get into certain honor societies, or get different types of recognitions or awards, which then impacts your selection into the Residency you want and your fellowship and downstream career opportunities, and it just compounds over time so that we get this kind of widening. And again, this amplification truly sad. And I've seen this play out multiple times, looking back on learners who all it takes is that slight edge to the better, or to your detriment, to really set you off on a different trajectory compared to someone else.
- But this isn't just potentially hurting our learners. It's also impacting our patients. Because while we might see the amplification cascade and just say, Okay, somebody's getting an unnecessary advantage. It's also about holding some people back in an unfair way, meaning the amplification. Cascade doesn't just push some people up. Inequitably it also pushes people down, and in some cases it means it hurts the diversity of our workforce downstream.
- And there's a growing body of evidence that things like racial concordance and gender concordance between doctors and patients improves communication and trust and adherence to medical advice. So if we are restricting the diversity of our workforce through inequitable assessment approaches, we are actually downstream, hindering our patients as well. So this is a moral issue. This is a learner harm issue. This is a patient harm issue and this is not a new issue. Like, I'm not saying anything. Novel people have been saying this for a long, long time and thankfully just. It seems to be getting a louder and louder chorus of voices around assessment, saying, We really really need to do better.
- And I want to point out that this is obviously not a simple problem. It is what Dr. Lucy calls a wicked problem. If you've not heard that term before. These are a wicked problem is something that does not have an easy solution. It may not have any solution at all. It might just be something that you try to mitigate and manage but it's filled with complexity and dynamic tensions and contextual influences and inequity in assessment is certainly that not the least of which, because, as I said earlier, we don't always agree on what we mean when we say equity or inequity of assessment. So it's really complex.
- But it's an issue that we need to address.

- And so what do we do? I don't want to just come and and problematize and and lame doom. I also want to say what are some things we can do about it, and I think some of the things that you're probably familiar with are really useful and important, like implicit bias, training and Dei training. I know all those things are like potentially under threat for political reasons, but keeping the discourse. Going upfront, I think, is a really important step. But there are other things we can do to specifically around assessment.
- One of the things that I would take us a long way is to give up on the idea of objectivity in assessment.
- Sometimes people try to improve equity and fairness of assessment by seeking objectivity. They'll say the problem is that some assessments are subjective, and those are the ones that are inequitable and some are objective, and those are the ones that are good.
- All assessment is subjective.
- In fact, John Norcini, who's somebody who tends to have a very like objective type of view, at least as far as I read his work, he just wrote this. This paper just came out like last month, and he wrote that all assessments are based on human judgments, regardless of whether they are considered subjective or objective.
- So those types of assessments that we usually hold as like objective assessments, because they come with numbers and confidence intervals, and it feels like everybody's getting the same thing like a multiple choice test or or an Osce. Those things still have human judgment baked into, how they were constructed, how the standards were set, and how they are deployed, there is still subjectivity baked into them, and human judgment baked into them, even if it's not in the same way as a radar based assessment that happens on the wards or in clinic. And so once we kind of let go of the idea of objectivity being the savior of inequity and assessment. I think we can move forward in other ways to try to mitigate these things. Otherwise I'm worried that we'll be chasing the pot of gold at the end of the rainbow that we're never going to catch, because no form of assessment is truly objective. All of it involves human judgment.
- I think another thing we can do is I mentioned this earlier. Take a very like process improvement. Approach to this, to try to fully understand the drivers of an equity assessment and take targeted interventions. Recognizing that there is no fix there is, I mean, this is going to be an ongoing iterative thing and that we only have finite resources to try to improve our assessment systems. So this is work again from Ucsf, Dr. Tehrani, where they were looking at their medical school, and they said, Let's really dive into the drivers and factors that are contributing to inequitable assessment, and they put them into 4 buckets, student factors, rater factors, learning environment factors, and then the policies and procedures of the school. And what I really love that they did other than the fact that they used a fishbone diagram is, they said, you know, some of these factors are experienced by all students kind of equally, and those things are in the dark black. So all students struggle with variable interest in a given course or learning environment. They come in with variable competency, variable levels of effort, and they all have performance, anxiety.
- But the things in the light gray on the right side, those are the things that they found were disproportionately, negatively impacting learners who come from marginalized or minoritized backgrounds, stereotype threat microaggressions, low social capital and belonging.

- And then you can start to really unpack. What are the things driving it, and have targeted interventions to try to improve these things? And they did the same thing for rater factors. The people who are actually giving the ratings for factors that are happening in the learning environment around assessment.
- And then the policies and procedures that are forming really the backbone of the assessment system. I thought this was a really clever way to approach it, and more frequently I think we should try to integrate quality improvement systems, improvement with education. I think this is a really great example of it and then you can see how it might unfold if you try to intervene. So this is a different group who used that approach to say, What if we actually start doing some interventions to fix these?
- Oh, shoot! Sorry, my yeah. Sorry. My slides got a little mixed up. But but they said, This is again, Catherine, Lucy they said, Okay, what are some interventions? We can do to try to intervene upon these these failure modes of these problems. And and many of them are things that align with competency, based assessment, kind of philosophies things like do criterion based assessment, not normative assessment. If you're not familiar with those terms, criterion based assessment says, don't have rankings and sortings. And you know who's 1st who's last. Set a criterion of what you're going for with your education in terms of like you know what is your line of pass fail or competent, and don't try to have rankings and sorting. People. Make sure that you have a lot of different assessment strategies. Don't put all your eggs in one basket in terms of one osce or one test. Use lots of different approaches to try to get a sense of what learners can do.
- Make sure your assessors are prepared because oftentimes, if you're doing observational assessment, the rater, the assessor is the assessment tool doesn't really matter what form you give them. It's about how well is the the assessor trained to do the job, and that includes things like implicit bias, training and training on equity.
- Make sure that you give people feedback informative assessment before you levy summative decisions.
- So people shouldn't be surprised or blindsided by getting a grade or an overall rating. If you haven't given them feedback ahead of time and given them a chance to improve make sure that your decisions are informed by performance in multiple contexts with multiple observers. Don't just say I saw you on rounds on, you know, this week on this service with this assessor, and we're going to draw big conclusions from that.
- Make sure, if you have a committee making decisions like a competency committee, that they have expertise in assessment and data analysis, and that you have a diversity of backgrounds, identities, experience all in that committee. So you get different viewpoints leading to the decisions.
- Make sure you're actually evaluating for ex equity in the processes and outcomes. And again, this is where understanding, what do you mean by equity is so important?
- Make sure everybody has fair access to learning opportunities, and then make sure you have kind of taking the temperature of your learning environment for all those things that can hinder performance may not get picked up in assessment, like stereotype threat, implicit bias, identity, safety, all of those things, and make sure that your learners have a mechanism for reporting when there are issues with those things.

- Now this is an example of a program that actually did some of these things. They followed that advice, and they had some really good outcomes. This is Dr. Eve Coulson at Washington University they said. We're going to improve access to our resources and learning experts so that people don't have to subsidize it for themselves. Everyone has equitable access to all the resources you would need to try to succeed here.
- It got rid of grades. It went to criterion based with their assessment system.
- They did a lot of training with their competency committee for bias mitigation.
- They improved their feedback, so people had opportunities to grow and and improve. Before getting a final assessment they really put a lot of effort into monitoring their learning, environment and improving their reporting processes for their trainees who might be experiencing adverse behaviors.
- And then they had a whole program evaluation going on that they put onto dashboards and regularly reviewed at high level committees to look for evidence of the process and outcome equity measures that were important to them. I thought this was really brilliant and just an example of how, you know, did they solve equity or fix equity and assessment? No, I think Dr. Colson would tell you that they still have issues and inequity. But this certainly was a significant step forward in improving things for their learners. Really brilliant. Actually we hadn't done any of that. What were we gonna do taking you back to? When Macy said, You've got a great assessment system. Can you write a paper about how equitable it was?
- We had never really like, held up our system and said, Is it equitable or not to to be to be honest with you, to be totally transparent and vulnerable? We had not, we had not done that. And so we we took the things that I just told you that Dr. Lucy put forward all those different ideas. Competency based approaches that can improve the equity of your system, at least from a process standpoint.
- And we found that we actually checked a lot of them. We were criterion based. I showed you we had lots of assessment strategies. We did a lot of faculty development. We had a whole coaching program with a dashboard for learners to see their data.
- We got lots of contexts and assessors and thousands of data points about our learners from lots of different opportunities to show what they could do.
- We spent a lot of time on recruiting a diverse Ccc. And making sure we had people with different types of expertise in it. And then I'm going to give this one a baby checkmark because we just started to look at equity in our processes and outcomes. We had just done a study on? Was there differences in entrustment, ratings based on gender identity of the assessor and of the learner? And whether you had gender discordance or concordance. We were just starting there, but we kind of checked a lot of these boxes.
- And we thought about writing that, but it still felt like, not totally great. We were like that. That's not. We weren't being very deliberate about equity. It would be. It would be disingenuous again for us to write that paper and so we were really struggling with why? And then
- I'm really grateful. Because Hannah Anderson, whose name I said earlier, we got a chance to to work with her and meet with her, and she really opened our eyes to something that I think was helpful, and that is that inequity is just kind of woven into the fabric of everything we do in assessment, which is why, no matter what type of assessment strategy you look at, you, find evidence of inequity. It's just part of how humans work.

- And this is a wonderful discourse analysis that her and her team did. And I'm just putting it up here because it drove the message home for me. Inequity is woven into the fabric of our assessment systems. And she said, You know, if you really want to hold a candle to whether your program is equitable or not, or how equitable it is. You got to ask some really difficult questions.
- Questions like, who is the system designed for? To begin with, who did you have in mind what was the learner you had in mind when you designed your assessment approaches? And who was there to help design it? Who wasn't there?
- Whose experiences and stories did you not hear? And are you still not hearing, and whose identities are you overlooking? And then, when you think about your processes which seem totally inert or impartial, who might you be actually hurting and marginalizing with that, and because you're not hearing their stories, you don't even know you're doing it.
- This, this really pushed us to some uncomfortable spaces, even though we were well intended. We had not been asking these questions to be honest with you. And so that was the paper we decided to write. We said, we're going to just actually tell the story of our journey, trying to get to those questions, hoping that the transparency would be helpful for other places. Some places are way ahead of us. Some places. Maybe you're are right with us, some behind us. But the journey of how we thought about assessment over that time. And we told the story through these 4 kind of little phrases that actually build on one another, starting with epistemology. And if you're not familiar with epistemology. It's essentially thinking about, how do we think we get to truth and knowledge in any given situation? So when you're thinking about assessment how do you? How do you think truth and knowledge exists, and how you get to it? About a learner's competence or readiness for practice, or something like that?
- And early in our assessment work at the beginning of the 15 years ago, when I talked to you when I mentioned we had a what's very called a very post positivist epistemology and post positivism, says there is an objective truth that you can know although you may never fully know it, because knowledge is always imperfect. There is an objective truth out there, and really, assessment tries to get at that truth about a learner. There's an objective truth about someone's competence. There's an objective truth about someone's readiness for practice all those kinds of things, and if only you had the best way to measure it then you could know it, and equity in that sense is about removing error from your measurement system. When we were doing all those learning analytics, and we were trying to account for all the you know, all those variables that we thought were introducing error into our numbers. This is the mind frame that we had. We were very post Positivist assessment was about measurement and equity required removing error from that measurement over time. We move to what's called a more constructivist epistemology. Constructivism, says there is no objective truth. Truth is socially constructed and created in a given moment, and I don't mean truth in terms of like, you know, is this coffee mug here, I mean, like truth about a learner's competence a social construct like that? Is there an objective truth about that? Or is everything socially created. Our understandings are socially created, and therefore assessment hinges on us, understanding the context in which that truth is created, and equity is more about co-creation and engagement with people who could help us co-create that understanding.

- And that's how our assessment system functioned for a long time, and what Hannah was pushing us toward was a more critical view of things, where you assume everything is shaped by power and that assessment inherently induces hierarchy and hegemony, and that you need to start critiquing the underlying structures and the history around what you've done. And this was really really helpful, because those 1st 2 orientations felt fairly inadequate and superficial, although they although they were helpful, they moved us in the right direction. But I really think this last jump is going to pay the most benefit to our learners and our patients, and what I didn't realize at the time, because she wrote her paper later is that we were really kind of journeying through the different orientations of equity. We started with the fairness, orientation that I described earlier, where it was about impartiality and removing error.
- Then we had journeyed into the inclusion section, where it was more about understanding context and co-creating with people and engagement with all learners and individuality. And then we were moving more toward a justice orientation.
- It's been a really interesting journey.
- So the last thing I'll just kind of finish on is what's on the horizon for improving equity and assessment, and you know I'll just be honest. I don't know that there is a solution again. I don't think this is a problem that can be solved. I think it's more of a wicked problem that we have to navigate and mitigate as we go. But some ideas that I think are happening just on the horizon of assessment that might move us forward somewhat.
- One of them is to start using clinical care measures in assessment, and the reason why I think this might be helpful is this type of assessment removes a human raider from the process.
- Humans are still involved in creating these assessment systems. So they still are prone to bias and inequity. But you wonder if we have other tools that might not rely on a human radar? Can we get more data that is somewhat different than the human radar based data that we so often rely on. So, for example, I mentioned this earlier, our residents get their clinical care measures.
- We track them over time while they're in their ambulatory long block. And what you're seeing here is a list of measures that every single one of our residents gets data on when you take over this panel of patients and you manage them for just over a year. What happens to these measures. And these measures, by the way, are actually chosen by our learners. Every year we have a conference called Defense of the Measures where the learners get to decide which measures do we take off this list? Which ones do we add? And what are the operational definitions that we're using to actually implement them?
- And we ask, can you hit the target performance which is either set by the learners or sometimes by the system? Or can you actually just nudge the measure better by, let's say, 2% over the year. And we just look to ask, can you do either one? Can you either hold the measure or can you improve it over time? This is one form of assessment, whether you use it strictly for giving feedback, or whether you put a summative bend to it that you can integrate into a system that again, would remove a direct observational radar from the process.
- Now, one thing you'll notice on here. Some of these measures may not be super highly attributable to an individual resident, and as an example, we broke these measures down along a continuum from attribution on the left side. These are

measures that are more attributable to the work a resident does down to contribution, which is something where there's so many factors going into it that the Resident only contributes one part of the outcome.

- And so some of those things are. Some of those measures are actions taken during the exam. So did you do a foot exam? Did you screen for depression? Did you talk about Psa screening?
- Some of those things happen while the patient's in the building, but they have to go down the hall to the lab to get blood drawn. So that's 1 step removed from the immediate ability of the Resident to impact some things actually require you to schedule a separate visit and come back like when you get your colon cancer screening. If you're doing colonoscopy or your breast cancer screening for mammography.
- And then some things are actual, more like outcome measures like your a 1 c and your blood pressure. And we recognize that, like, you know, some of these measures might be better than others for actual use and assessment. And so the next iteration of this is something that is being led by Dr. Jesse Berkoffel at Nyu, called tracers. And this is essentially where we are engineering clinical measures and choosing measures that are specifically useful for assessment, because they are more attributable to a learner, and then they have all these other characteristics that you see there so measures that are meaningful for care attributable to an individual trainee, not 100 attributable, but attributable enough to be meaningful to the residents of the program, automatable, scalable in real time.
- And just as an example. After talking with residents and faculty at our 2 institutions, we developed measures around ordering of insulin for patients admitted who have type, one or type, 2 diabetes.
- And what you're looking at here is each one of these dots represents a different resident at Nyu. So at Jesse's home program and on the X-axis is the rate at which they order short acting insulin for any patient admitted with diabetes, and you'll see some of the dots which are again. Different residents are all the way over to the right.
- Those are people who order short acting, insulin 100% of the time every time they admit somebody with diabetes on the Y-axis is the rate at which you order long acting insulin. And so the people who are on the bottom of this graph order, long acting, insulin, 0% of the time.
- The orange dots that are on that line in the middle are people who order them at the exact same rate, so every time they order one they order the other as well and for each one of these dots you can't say whether the ratio with which they ordered them is right or wrong, because there's so many like unique, patient factors. But you could start pairing this with the glycemic control numbers and saying, You know, hey, this person? They order short acting, insulin all the time, but never order long acting insulin. If their glycemic control numbers are terrible. Maybe they need to start integrating long acting insulin into their practice patterns and their practice patterns are probably very different from this person, which is different from this person there might be different phenotypes of learners in there. And we can start using this type of measure for feedback, and it does not rely on a human rater to get you that data.
- We've also started using metadata in how people are using the Ehr so simple things like seeing how long people spend in notes, and what's the length of their documentation? And what's the turnaround time on the orders that they that they

put in, or the refills that they do again. None of this is to set a bar or be big brothery. But to say, if you have somebody really struggling with efficiency, this is one form of assessment that can be helpful for them, and does not rely on human raters.

- Now, again, all of these could have issues with bias and inequity, because all forms of assessment have human judgment. But I think the more diverse sets of assessment data we have, the more holistic of a judgment we can make about people, whether it's used to coach and get them better, or whether it's used to make summative decisions.
- And lastly, I'll just say, AI
- AI is fraught I think it's well known, and it's a very well known discourse. The AI brings lots of risk for bias and inequity, just like anything else.
- But I think there's also opportunities to integrate that into assessment. As long as we are aware of the biases and inequities that come with it, we're vigilant about them, and we try to mitigate them as they come up. I think it is just as risky as human raiders, but it is again another tool that we can gather data. And this is work from Dr. Verity Shea out at Nyu, where they are using AI to give feedback to learners on the differential diagnoses that they put in their notes. How's their explanation of clinical reasoning and giving learners a real time dashboard to try to reflect on? How are they expressing their clinical reasoning and their documentation and kind of even.
- And this is another example of AI. This is Dr. Laura Turner and Seth overlay here in Cincinnati, using their 2 Sigma platform to do real time dynamic cases with our medical students and giving them feedback on their clinical reasoning and their medical knowledge, and we haven't used it for summative judgments at all. But it's a really powerful formative assessment tool that, again, is free of a human judgment. There is human, free of a human raider. There's human judgment in building this, and they are very actively working on and monitoring for issues with equity and bias, including the outcomes of its use. So it's not free of those things. But it's just another kind of arrow in the quiver and the last thing I'll say is I think that technology is going to take us even farther with things like haptics.
- You're not familiar with haptics. Haptics are essentially motion capture devices that can give us feedback on all kinds of things. And what you're looking at here is laparoscopic surgery, where you can wear motion tracking sensors and actually track the even the most minute movements that somebody does during a surgery figure out which movements and phenotypes of surgical approaches are most are most associated with better outcomes. And you can start developing these kind of fingerprints of how people practice and giving people feedback or using them even summatively, to give as part of an assessment program.
- And even if you're not a procedural specialty, there's all sorts of other types of haptics like eye motion capture technology where you can look at, how do people approach a worried parent or a sick infant? And where do their eyes go when they're in their communication? Or how do they go about navigating through an Ekg or a chest? X-ray? Or how do they go about assessing a rash or a skin lesion? All of these things again we're still just the very beginning of learning how to understand them. But all of them provide a potential step in the right direction around not relying as much on human raters, not in trying to get to objectivity, because it's not possible, but to diversify our assessment approaches in an attempt to try to to be more equitable.
- I think all these things are on the horizon and potentially, very, very useful.

- But I'll I'm just gonna close with this, that the conversation around equity, both just in assessment, but more broadly and in our society, is changing really, really fast.
- I think that one of the things we can do is be open to the idea that the way we think about equity now, or the way we thought about in the past may not be how it is in the future. And one of my favorite quotes from All time comes from this editorial by Kevin Eva. It's a short, short editorial. It's 2 pages, but he writes that health professions, education would be an uninspiring place if every idea put forward was guaranteed to be long lasting.
- Think I think we need to bring a lot of humility to these discussions that our understandings change the learners of our needs and our the needs of our learners and our patients change very quickly, and we need to be flexible and adaptable and open to the fact. Things are going to change and probably the most important thing I'll say today, and it's the last thing I'll say is that as we do this work, most important thing is the mindset we bring to it.
- I think we need to bring curiosity to seek, 1st to understand and then to be understood.
- As I said, humility to admit that our assessment systems are flawed, and that we don't have the answers to fix them right away.
- And we bring grace to allow for mistakes and failures as we do this work because we will all make mistakes, and we will all have failures that we have grit to persist in this work, despite social or political headwinds that we have urgency to try to make changes because all of our learners are being impacted by this amplification cascade, and our patients are being impacted as well.
- And then, lastly, joy, this is a huge opportunity to make a profound impact on the on the lives and careers of our learners and on our patients, and so we should have, we should bring a spirit of joy to this. I feel like a lot of the a lot of the conversation can be very dour and sad because it is such a big problem. But anytime there's a big problem. There's a big opportunity to improve someone's life.
- If anything that I said is interesting to you, and you want to hear someone smarter and more well spoken. Talk about it, or write about it. I highly recommend just plugging any one of these names into pubmed or Google scholar. And you'll you'll be hit with a tsunami of wisdom. These are all amazing people that I look up to, and I love the things that they do and the work that they do cannot recommend them strongly enough.
- And hopefully I represented some of their ideas fairly well here, and I am so excited to hear any reflections people have, or ideas, and thank you very much for the invitation to come join me today.

Thank you, Dr. Kinnear. That was excellent. I will open up the questions with my own question, so I wanted you to expand a little bit more on what you guys are doing in your program where the residents are like selecting for the different criteria and their ambulatory setting for like improvement in patient care, and also like with respect to likenin internal medicine, you have residents of who have a variety of interests, some who aren't necessarily interested in primary care and maybe are more interested in the inpatient setting. And you know, steps that you guys are taking or thinking about to consider those residents with different interest, or are going to be primed to be thinking clinically, differently about a patient that they approach.

Yeah. So to your first, st to your 1st question. The the like. I said. The thing that we call that moment of residents engaging in the co-creation of their of their performance. Measure dashboard. We call that defense of the measures happens once a year. It's a big event.

- We do it. During our academic half day all residents are invited, all of our clinic staff is invited as well as our faculty. Everyone's voice and vote counts the same. We set it up debate style, where the residents choose which measures are on there, and the evidence has changed, and they think it's time to change the operational definition, or they think it's time to boot that measure off there. Or maybe new evidence has come in and they want to add something new.
- They have a brief debate, and then we vote, we discuss and we vote. And I think there's multiple benefits to that. One is giving agency to the learners about the things and the ways that they're being assessed, I think, is really important, both in terms of getting their buy-in on the assessment process, but also it has an element of equity and fairness to it. Again, like I, said, one of the orientations to equity is about co-creation with learners, and and that's a big part of it.
- As far as the inpatient. You know there have been a couple of learners who have opted out of long block for one reason or another, and we're not totally rigid about that. But during recruitment for the program, Dr. Warm, the medicine program doctor is very upfront about long block. And the goal of the program is not to create only primary care providers, but to give people a better sense of what does it mean to do primary care? Because even if you're going to be a critical care, Doc, or cardiologist, or a hospitalist having a good understanding of what primary care is and how it works is beneficial for your career. And I think you know I'm the Med Peds program director. We don't have long block in my program. I think our residents have a different understanding of what it means to be a primary care, doctor, than our long block residents who spend an entire year embedded in there, truly doing the work of a Pcp.
- I think it's I think it's incredibly beneficial, regardless of what? What program you're going into. But we you know Eric is very upfront about that during recruitment to say, if spending an entire year as a primary care doctor doesn't sound like fun to you. You probably shouldn't come to our program.

We do have a couple of other comments in the chat, talking about great talk, as well as from Dr. Luna, comprehensive and systematic analysis to the challenges and opportunities to improve our learners career and our relationship with them. And then, Dr. Mamari, thanking you for the talk and giving residents, the autonomy to determine their own outcomes is fascinating and exciting. Would you attempt a similar approach and assessment of clerkship learners?

- In other words, should they choose how much osce versus observations versus shelf exams count. Oh, that's such a great question. Great question, Dr. Mamari. I would say. Maybe
- I think it depends a lot on
- It depends a lot on what other assessment strategies you have, because I think good assessment and equitable assessment is a program of assessment where you have multiple different approaches. And so in some parts of that program I

would give learners autonomy to to figure out. What are they going to be assessed on? When are they going to be assessed?

- Especially because if you don't have any spaces for assessment that is not linked to downstream judgments, in other words, assessment, just for the sake of feedback. You don't have any space for that. It's really hard for learners who are under immense pressure to really be vulnerable and take chances in their learning, and to grow so if you give them some agency about to choose their choose their spots. When is it when it when, and how are they going to be assessed? I think that's helpful. That being said you know, there is this concept in in assessment and measurement of reliability? How many times you need to see something happen before you can say. I think I understand what's going on here, and it's not just random chance there may be some parts of your assessment system where you need to have a certain number of observations or measurements, or whatever you want to call it, for you to feel good about the judgments that you're going to make downstream. And in those cases maybe you can't give them total agency, because maybe they would never get to the number that you need. So I think it's all very dependent. But the more agency you can put in some parts of your assessment system. I think that contributes to the overall process equity of what you're doing. And I think it really is beneficial for learners in terms of their investment in what's happening with them. Excellent, and then we'll probably this will be our final question. This is from Marcus, one of my co-chiefs. So thank you. Very interesting concepts in your work. Have you or colleagues encountered any organizational or institutional barriers to pushing this work forward? What do you see as helpful for weathering the storm of pushback to Dei work. Oh, man, great question I I wondered if something like this would come up.
- This is where I'll caveat my response with the fact that I am not an equity scholar. I'm not an advocacy expert. I've been very lucky that our institutions here, even though when you hear Ohio, you might not think.
- Dei, you know, a bastion of Dei, but I think many academic centers probably recognize the value of and importance of equity and diversity, and belonging and inclusion.
- I'm lucky that my institutions recognize that value, and they are working really hard to follow our missions that include equity and inclusion and diversity, even though there are political, specifically political headwinds in in doing that work.
- How?
- What is helpful in weathering the storm?
- This is just my opinion. This is not even space. There are probably much more intelligent people on this call who could? Who could speak to it? I think community is huge.
- I think trying to weather this by yourself is really really hard, and finding communities of people to help you move forward is helpful.
- I also think hope is really huge.
- There's a quote, and I'm sure I'm gonna I'm gonna butcher. The quote. But hope is not a lottery ticket that you sit on the couch waiting for a good outcome. Hope is an axe that you grab in an emergency and use it to break down doors. I think that having hope will lead you to action as a community and as a group will lead to change.

- And I think if you don't link the things you believe in. You think equity is important. If you don't link the things you believe into action, then you're never going to have change.
- That's what I would say.

UVA Grand Rounds

01:30:48

Thank you so much, Dr. Kinnear. This was a great talk. Appreciate your time.

Benjamin Kinnear (he/him)

Thanks for the invitation. Everybody!