

Did You Catch These Non-JGME Medical Education Articles From 2022? Editors' Recommendations

Gail M. Sullivan, MD, MPH

Deborah Simpson¹, PhD

Anthony R. Artino Jr¹, PhD

Lalena M. Yarris¹, MD, MCR

Nicole M. Deiorio¹, MD

For several years the *Journal of Graduate Medical Education* (JGME) senior editors have assembled articles published in other health professions education publications in the past year that we think are worthy of notice. As we have different interests, areas of expertise, and definitions of usefulness, our conversations are animated and the resulting collection is eclectic. There is no scientific approach to this process; we simply like these articles and hope they may prove helpful to you, too (see TABLES 1 and 2). Tell us what you think by tagging @JournalofGME on Twitter.

Tony Artino's Pick

The last time you took a survey, whether it was about faculty satisfaction or an evaluation of your hotel stay, did the survey ask you to agree or disagree with a set of statements? I would be willing to bet it did, since survey items with agree-disagree response options are the most commonly used format to assess attitudes and opinions. For example, in a medical education study I led a few years ago, we found that 57% of published surveys included at least one agree-disagree item, and across all of the survey items reviewed in our sample, 45% of items used agree-disagree response categories.¹ The ubiquity of agree-disagree items is not surprising, because they are easy to write—simply create a list of statements and then ask respondents to agree or disagree with those statements. However, the utility of such items and their psychometric properties have long been debated. In this review by Dykema and colleagues, *Towards a Reconsideration of the Use of Agree-Disagree Questions in Measuring Subjective Evaluations*,² the authors focus on the measurement properties and potential limitations of agree-disagree items compared to what they term *item-specific questions*.

An example of an agree-disagree item compared to a corresponding item-specific question is:

- *Agree-disagree*: To what extent do you agree or disagree with the following statement? I feel well prepared to perform laparoscopic surgery without supervision. (Strongly disagree, disagree, neutral, agree, strongly agree)
- *Item-specific*: How well prepared are you to perform laparoscopic survey without supervision? (Not at all well prepared, slightly well prepared, moderately well prepared, quite well prepared, extremely well prepared)

Agree-disagree items present respondents with a statement and then ask them to rate their level of agreement, whereas item-specific questions directly ask respondents about the underlying construct being assessed (in this case, *perceived preparedness*) using response categories tailored to match the construct.

Dykema and colleagues reviewed 20 experimental studies that directly compared agree-disagree and item-specific questions.² Although mixed, the findings indicate that most studies reporting item-specific questions are associated with greater reliability and validity when compared to agree-disagree items. The authors note several explanations for these results, which correspond to what we know about how respondents work through the cognitive steps needed to answer survey questions. The authors' reasons for survey designers to avoid agree-disagree items include: respondents are more likely to acquiesce (ie, agree) when answering agree-disagree items compared to item-specific questions; agree-disagree items often present respondents with a mismatch between the item's underlying response dimensions and the response options offered by the agree-disagree categories, which adds to cognitive burden; and agree-disagree items use bipolar response options that present both ends of a negative to positive response

DOI: <http://dx.doi.org/10.4300/JGME-D-22-00959.1>

TABLE 1
Noteworthy Non-JGME Articles From 2022

Towards a reconsideration of the use of agree-disagree questions in measuring subjective evaluations	Dykema J, Schaeffer NC, Garbarski D, et al. <i>Res Social Adm Pharm.</i> 2022;18(2):2335-2344. doi:10.1016/j.sapharm.2021.06.014
The otolaryngology residency program preference signaling experience	Pletcher SD, Chang CWD, Thorne MC, et al. <i>Acad Med.</i> 2022;97(5):664-668. doi:10.1097/ACM.0000000000004441
Climate change: a crisis for family medicine educators	DeMasi M, Chekuri B, Paladine H, et al. <i>Fam Med.</i> 2022;54(9):683-687. doi:10.22454/FamMed.2022.827476
Self-assessment: with all its limitations, why are we still measuring and teaching it? Lessons from a scoping review	Yates N, Gough, S, Brazil V. <i>Med Teach.</i> 2022;44(11):1296-1302. doi:10.1080/0142159X.2022.2093704.
Burnout, wellbeing and how they relate: a qualitative study in general practice trainees	Prentice S, Elliott T, Dorstyn D, Benson J. [Published online ahead of print August 23, 2022]. <i>Med Educ.</i> doi:10.0000/medu.14931

dimension, whereas item-specific questions can be either bipolar or unipolar.

My take-home: Although experimental studies directly comparing agree-disagree and item-specific question yielded mixed results, more studies found that item-specific questions are associated with desirable data and that agree-disagree items are associated with undesirable data quality. Therefore, I—like the authors of this excellent article—recommend item-specific questions over agree-disagree items for *most* survey purposes. Read the article for yourself and make your own evidence-informed decisions the next time you are designing a graduate medical education (GME) survey for research or evaluation.

Nicole Deiorio's Pick

Imagine a world where residency applicants match to a GME program they love, after spending a reasonable amount of money on applications, and are welcomed enthusiastically by the program director as a top choice applicant.

Not possible? We edge closer to this dream as we gather more data around recent Match innovations conducted in the last couple of application cycles. In *The Otolaryngology Residency Program Preference Signaling Experience*, Pletcher et al describe the first year of the otolaryngology preference signaling trial, in which applicants could signal up to 5 programs at the time of initial application as an indication of special interest in the program.³ In this relatively small competitive specialty, 558 of the total 559 applicants employed the signaling process. Surveyed program directors (52% response rate) reported that the rate of receiving an interview offer was higher from signaled programs (58%) than from non-signal programs (14%; $P < .001$) and the next non-

signaled program (23%; $P < .001$; ie, the program an applicant would have signaled given a sixth signal). Interestingly, these differences were seen across the range of applicant competitiveness. Surveyed applicants (42% response rate) and program directors strongly favored continuing the program.

My take-home: This article lays the groundwork for essential outcomes-based research for all the Match innovations from the current and future cycles.

Deb Simpson's Pick

In late July 2022 Wisconsin sunrises and sunsets appeared a beautiful yet unusual hue of red. While breathtaking, the cause was smoke from fires on the West Coast and Canada. In September 2021, more than 200 medical journals published a joint editorial that called climate change “the greatest threat to global public health.”⁴ Medical education journals have published articles that address training for mass trauma, wildfires, extreme weather events, pandemics, and vector-borne diseases and often link these phenomena to climate change. Currently, 15% of medical schools worldwide teach a climate and health curriculum,⁵ with a group of students leading an international planetary health report card.⁶ There are limited reports around GME⁷ and no student or resident accreditation requirements on this topic.

In 2022, *Family Medicine* published a commentary by DeMasi and colleagues, *Climate Change: A Crisis for Family Medicine Educators*, in which the authors issued a similar call to action.⁸ Yet, this one stands out. After outlining the inequities of climate change on patients and adverse effects on health care clinicians, the authors describe why climate change is “in our lane” as educators, include resources with specifics for GME, and suggest actions we can take. Normally, calls to action don't actually push me into

TABLE 2

Honorable Mentions, Non-JGME Articles From 2022

<p>An outcomes-oriented approach to residency selection: implementing novel processes to align residency programs and applicants Caretta-Weyer HA. <i>Acad Med.</i> 2022;97(5):626-630. doi:10.1097/ACM.0000000000004614</p>	<p>This article is a commentary on our “notable” paper by Pletcher et al on the otolaryngology residency program signaling initiative. Dr. Caretta-Weyer writes that the current proposals to address application inflation, which focus on logistics (application caps, implementing an early Match, signaling) will not address “root causes,” such as medical students’ fears, and may perpetuate inequities. An “outcomes-oriented selection process” and residency program selection transparency are required.</p>
<p>Diversity, equity, and inclusion competencies across the learning continuum Association of American Medical Colleges. New and Emerging Areas in Medicine Series. July 2022. https://www.aamc.org/data-reports/report/diversity-equity-and-inclusion-competencies-across-learning-continuum</p>	<p>This publication specifies DEI competencies as a continuum with 3 stages: entering residency, entering practice, or faculty physician teaching and leading. The continuum is framed using the 3 domains of DEI. Each domain has sequenced competencies, 24 in total. This is a go-to DEI reference useful across the medical education continuum.</p>
<p>The association between USMLE Step 2 clinical knowledge scores and residency performance: a systematic review and meta-analysis Shirkhodaie C, Avila S, Seidel H, et al. <i>Acad Med.</i> [Published online ahead of print November 1, 2022]. doi:10.1097/ACM.0000000000005061</p>	<p>This systematic review and meta-analysis used strong methods to search databases, 1992 to May 2021, for studies comparing USMLE Step 2 CK results with resident performance, including ITE scores, board certification scores, overall resident performance, and other performance measures related to ACGME competencies. The authors examined surgical and non-surgical fields independently as well. Of 68 included studies, 43 could be pooled. As expected, there was moderate correlation with Step 2 CK and ITE scores—good test-takers stay good test-takers—and a very weak correlation (0.19, 95% CI 0.13-0.25, $P < .01$) for resident performance assessed by individuals. This is likely the best summary of data on this question, and it should point program directors away from using Step 2 CK as a new applicant filter.</p>
<p>Virtual interviewing for graduate medical education recruitment and selection: A BEME systematic review: BEME Guide No. 80 Daniel M, Gottlieb M, Wooten D, et al. <i>Med Teach.</i> 2022;44(12):1313-1331. doi:10.1080/0142159X.2022.2130038</p>	<p>The authors scoured the literature, from 2012 to February 2022, and included applicants to residency and fellowship. Although nearly all the 110 included papers were from North America, the findings—time and money saved, less confidence in ranking, and applicant virtual interview preferences—can be helpful to those implementing and studying virtual interviewing, which appears here to stay.</p>
<p>Crossing the virtual chasm: practical considerations for rethinking curriculum, competency, and culture in the virtual care era Bolster MB, Chandra S, Demaerschalk BM, et al. <i>Acad Med.</i> 2022;97(6):839-846. doi:10.1097/ACM.0000000000004660</p>	<p>At the start of the COVID-19 pandemic, patient care rapidly transformed from in-person to virtual patient care, with little prior evidence to direct medical education in this setting. This article summarizes recommendations from a symposium of US leaders held in September 2020 around educational tools for training in virtual patient care: core competencies, assessment tools, precepting workflows, and technology. These recommendations, based on experts and some pre-pandemic telehealth studies, are highly practical.</p>

Abbreviations: DEI, diversity, equity, and inclusion; USMLE, United States Medical Licensing Examination; CK, clinical knowledge; ITE, in-training examination; ACGME, Accreditation Council for Graduate Medical Education.

action. This one did, because links to the article were sent to me from individuals in multiple GME programs locally and nationally.

My take-home: This commentary is a short read that has prompted us to take real, climate-related, education actions.

Gail Sullivan’s Pick

I was drawn to the article, *Self-Assessment: With All Its Limitations, Why Are We Still Measuring and Teaching It? Lessons From a Scoping Review*, by Yates and colleagues, because we use self-assessment ubiquitously for trainees and faculty.⁹ We do this

TABLE 3
When and When Not to Use Self-Assessment

Use	Don't Use
<ul style="list-style-type: none"> • Measure subjective outcomes: confidence, anxiety, empathy • Guided self-assessment (ie, avoid unguided assessment without focusing the trainee on clear, specific standards to aim for) • Study self-assessment as a process, not an ability 	<ul style="list-style-type: none"> • Measure trainee performance, competence, knowledge, skills • Evaluate programs or interventions • Study the accuracy of self-assessment, with goals of improving performance or lifelong learning

From: Shirkhodaie C, Avila S, Seidel H, et al. The association between USMLE Step 2 clinical knowledge scores and residency performance: a systematic review and meta-analysis. *Acad Med*. [Published online ahead of print November 1, 2022]. doi:10.1997/ACM.0000000000005061

despite research showing that external assessment is superior, for physicians in particular, because there is poor correlation of self-assessed learning with measures of competence.¹⁰ The authors' methods were also intriguing: from their original search, more than twice as many studies used self-assessment inappropriately vs appropriately, which piqued the authors' curiosity. A special treat in this article is the authors' self-assessment example—teaching a teenager to drive a car. Read this article for the vignette alone!

The authors examined the extent of self-assessment use for medical students in non-evidence-based ways, which was defined as it being a sole or primary outcome measure for assessing a program or intervention, or for self-assessment as a learning goal itself (ie, accuracy of self-assessment), as studies show it is not associated with improving performance or lifelong learning. The authors consider overlapping concepts, such as self-evaluation, self-monitoring, and self-efficacy, which may obscure studies of self-assessment. They found that in 63 of the 207 articles (30%), self-assessment of knowledge or skills was the sole outcome measure for evaluating a program or intervention. In 62 studies (30%), self-assessment of confidence was measured: when confidence and competence were both measured, correlation was variable, as found in prior studies. In 39 studies (19%), the study aim was limited to furthering the accuracy of self-assessment.

The authors were guided by Arksey and O'Malley's framework for scoping reviews, their protocol was registered on the Open Science Framework, and they used strong methods throughout.¹¹

Despite the focus on medical students in this article, there is abundant overlap with assessment methods used in GME. Many programs routinely measure self-assessed performance without additional measures, and JGME receives many papers in which self-assessed knowledge or skills are the sole outcome measure. It's worrisome that many articles in this review were published in "top tier" medical education journals and that there was no decline in the number of articles over the years of the review.

My take-home: Self-assessments of knowledge, skills, or confidence are rarely helpful, despite being easier to collect. Let's choose better outcomes *and* reduce our overall data collection load at the same time (TABLE 3).

Lainie Yarris's Pick

Wellness and burnout continue to be hot topics in GME, yet most studies are descriptive in nature. We know that burnout is an ongoing problem: we have measured its prevalence in many specialties, and we are starting to see studies that either seek to explore factors related to burnout or evaluate interventions aimed at improving burnout. However, deep understandings of the nature of burnout and well-being are still elusive, which limits our ability to design and implement changes that result in a meaningful difference in trainees' experience of burnout. Vexing questions abound: Why do some trainees experience tremendous stress, but not burnout, while others struggle? What trainee, program, and culture factors are associated with resilience and burnout recovery? How do personal characteristics, program factors, workload, and other stressors interact with the experience of well-being?

In this qualitative study, *Burnout, Wellbeing and How They Relate: A Qualitative Study in General Practice Trainees*, Prentice et al apply a post-positivist epistemology and grounded theory approach to explore the concepts of well-being and burnout from the perspectives of Australian general practice trainees and registrars.¹² Their subjects describe burnout as a syndrome that exists on a spectrum. Both trainees and registrars identified 7 relevant themes: altered emotion, compromised performance, disengagement, dissatisfaction, exhaustion, overexertion, and feeling overwhelmed. Well-being involved a complex interaction between factors in personal and professional domains, with an underlying "well-being reservoir" as an important facilitator of the perception of wellness. The authors propose an overarching explanatory model that centers around the observation that burnout occurs when a trainee's values and/or goals are no longer met, thereby depleting the well-being reservoir. The model

proposes that unfulfillment of an individual's professional and personal values is the central process by which burnout develops, which has important implications for the role of values in future interventions to promote well-being and address burnout.

This study caught my eye for several reasons. First, because it is a rigorous qualitative exploration of important phenomena. It adheres to published standards for quality and rigor in qualitative work¹³ and has strong rationales for the research paradigm, approach, data collection, and analysis methods, as well as appropriate justification for controversial methodological decisions. Also, the article takes a step back to question the status quo. Rather than continuing to build off prior models and apply existing tools to inform the development of interventions, the researchers ask an important question: *Do we really understand well-being and burnout in terms of how GME trainees experience these phenomena?* Finally, this article is timely and relevant. The past 3 years have been *hard*. For trainees, for educators, for our loved ones, for our patients. Not a day goes by where the topics of well-being and burnout are not front and center in at least one of my conversations with friends, family, learners, or colleagues.

My take-home: The concepts and findings in this article resonate deeply, and the proposed model is a thought-provoking and worthwhile read for medical educators and leaders.

References

1. Artino AR Jr, Phillips AW, Utrankar A, Ta AQ, Durning SJ. "The questions shape the answers": assessing the quality of published survey instruments in health professions education research. *Acad Med.* 2018;93(3):456-463. doi:10.1097/ACM.0000000000002002
2. Dykema J, Schaeffer NC, Garbarski D, et al. Towards a reconsideration of the use of agree-disagree questions in measuring subjective evaluations. *Res Social Adm Pharm.* 2022;18(2):2335-2344. doi:10.1016/j.sapharm.2021.06.014
3. Pletcher SD, Chang CWD, Thorne MC, et al. The otolaryngology residency program preference signaling experience. *Acad Med.* 2022;97(5):664-668. doi:10.1097/ACM.0000000000004441
4. Full list of authors and signatories to climate emergency editorial September 2021. *BMJ.* Accessed November 28, 2022. <https://www.bmj.com/content/full-list-authors-and-signatories-climate-emergency-editorial-september-2021>
5. Omrani OE, Dafallah A, Castillo BP, et al. Envisioning planetary health in every medical curriculum: an international medical student organization's perspective. *Med Teach.* 2020;42(10):1107-1111. doi:10.1080/0142159X.2020.1796949
6. Planetary Health Report Card. Accessed November 28, 2022. <https://phreportcard.org/>
7. Philipsborn RP, Sheffield P, White A, Osta A, Anderson MS, Bernstein A. Climate change and the practice of medicine: essentials for resident education. *Acad Med.* 2021;96(3):355-367. doi:10.1097/ACM.0000000000003719
8. DeMasi M, Chekuri B, Paladine H, et al. Climate change: a crisis for family medicine educators. *Fam Med.* 2022;54(9):683-687. doi:10.22454/FamMed.2022.827476
9. Yates N, Gough, S, Brazil V. Self-assessment: with all its limitations, why are we still measuring and teaching it? Lessons from a scoping review. *Med Teach.* 2022;44(11):1296-1302. doi:10.1080/0142159X.2022.2093704.
10. Eva K, Regehr G, Gruppen LD. Blinded by "insight": self-assessment and its role in performance improvement. In: Brian David H, Lorelie L, eds. *The Question of Competence: Reconsidering Medical Education in the Twenty-First Century.* Cornell University Press; 2012:131-154.
11. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8(1):19-32. doi:10.1080/1364557032000119616
12. Prentice S, Elliott T, Dorstyn D, Benson J. Burnout, wellbeing and how they relate: a qualitative study in general practice trainees. *Med Educ.* [Published online ahead of print August 23, 2022]. doi:10.0000/medu.14931
13. O'Brien BC, Harris IB, Beckman TJ, et al. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med.* 2014;89(9):1245-1251. doi:10.1097/ACM.0000000000000388



Gail M. Sullivan, MD, MPH, is Editor-in-Chief, *Journal of Graduate Medical Education* (JGME), and Associate Director for Education, Center on Aging, and Professor of Medicine, University of Connecticut Health Center; **Deborah Simpson, PhD**, is Deputy Editor, JGME, and Director of Education, Academic Affairs at Advocate Aurora Health, and Clinical Adjunct Professor of Family & Community Medicine, Medical College of Wisconsin, University of Wisconsin School of Medicine and Public Health; **Anthony R. Artino Jr, PhD**, is Deputy Editor, JGME, and Professor and Associate Dean for Evaluation and Educational Research, The George Washington University School of Medicine and Health Sciences; **Lalena M. Yarris, MD, MCR**, is Deputy Editor, JGME, and Professor of Emergency Medicine, Oregon Health & Science University; and **Nicole M. Deiorio, MD**, is Executive Editor, JGME, and Professor, Department of Emergency Medicine, and Associate Dean, Student Affairs, Virginia Commonwealth University School of Medicine.

Corresponding author: Gail M. Sullivan, MD, MPH, University of Connecticut Health Center, gsullivan@uchc.edu, Twitter @DrMedEd_itor