



The Next Era of Assessment: Building a Trustworthy Assessment System

COLLECTION:
NEXT ERA IN
ASSESSMENT

EYE OPENER

HOLLY A. CARETTA-WEYER

ALINA SMIRNOVA

MICHAEL A. BARONE

JASON R. FRANK

TINA HERNANDEZ-BOUSSARD

DANA LEVINSON

KIKI M. J. M. H. LOMBARTS

KIMBERLY D. LOMIS

ABIGAIL MARTINI

DANIEL J. SCHUMACHER

DAVID A. TURNER

ABIGAIL SCHUH

*Author affiliations can be found in the back matter of this article

ubiquity press

ABSTRACT

Assessment in medical education has evolved through a sequence of eras each centering on distinct views and values. These eras include measurement (e.g., knowledge exams, objective structured clinical examinations), then judgments (e.g., workplace-based assessments, entrustable professional activities), and most recently systems or programmatic assessment, where over time multiple types and sources of data are collected and combined by competency committees to ensure individual learners are ready to progress to the next stage in their training. Significantly less attention has been paid to the social context of assessment, which has led to an overall erosion of trust in assessment by a variety of stakeholders including learners and frontline assessors.

To meaningfully move forward, the authors assert that the reestablishment of trust should be foundational to the next era of assessment. In our actions and interventions, it is imperative that medical education leaders address and build trust in assessment at a systems level. To that end, the authors first review tenets on the social contextualization of assessment and its linkage to trust and discuss consequences should the current state of low trust continue. The authors then posit that trusting and trustworthy relationships can exist at individual as well as organizational and systems levels. Finally, the authors propose a framework to build trust at multiple levels in a future assessment system; one that invites and supports professional and human growth and has the potential to position assessment as a fundamental component of renegotiating the social contract between medical education and the health of the public.

CORRESPONDING AUTHOR:

**Holly A. Caretta-Weyer, MD,
MHPE**

Department of Emergency
Medicine, Stanford University
School of Medicine, 900 Welch
Road, Suite 350, Palo Alto, CA
94304, US

hcweyer@stanford.edu

TO CITE THIS ARTICLE:

Caretta-Weyer HA, Smirnova A, Barone MA, Frank JR, Hernandez-Boussard T, Levinson D, Lombarts KMJM, Lomis KD, Martini A, Schumacher DJ, Turner DA, Schuh A. The Next Era of Assessment: Building a Trustworthy Assessment System. *Perspectives on Medical Education*. 2024; 13(1): 12–23. DOI: <https://doi.org/10.5334/pme.1110>

Assessment systems are socially constructed and contextually embedded. As such, assessment criteria reflect current mental models of competence, both explicit and implicit. Assessment methods often reflect accreditation requirements and local practices which are not necessarily best practices. Human beings engaging in assessment bring biases and variable attention and abilities to the enterprise. Assessment criteria and processes subsequently reflect contemporary criteria for competence, values, and approaches to decision making. Much of assessment's evolution to date has focused more on the technical aspects of measurement and implementation of assessment, with less attention paid to the social context of assessment which we posit has led to an overall erosion of trust in assessment [1]. This lack of attention to assessment social context represents a failure of rapport, inattention to the contextualization of assessment, and a paucity of direct feedback for continued growth and development between all stakeholders across all levels ranging from micro to global. Multiple industries and professions use assessment to demonstrate accountability to the populations they serve, by ensuring individuals demonstrate the abilities needed to meet the needs of those populations [2, 3]. As assessment in medical education has evolved over time, these factors have contributed to a degradation of trust between individuals, programs, and systems [4, 5]. In order to meaningfully move forward, we must seek to reestablish trust in assessment at all levels within medical education.

Assessment in medical education has evolved through a sequence of eras each centering on distinct views and values, as described eloquently by Schuwirth and van der Vleuten [6]. These eras have included measurement (e.g., knowledge exams, objective structured clinical examinations), then judgments (e.g., workplace-based assessments, entrustable professional activities) and most recently systems, or programmatic assessment, where over time multiple types and sources of data are collected and combined by competency committees to ensure individual learners are ready to progress to the next stage in their training. Ideally, the continued evolution of assessment would progress in pursuit of public and professional accountability by building on the advantages of the previous eras, preserving what is good and addressing what needs to change.

To be accountable to the public, assessment in medical education must build and maintain the trust of both learners (in that they are being fairly assessed) and society [7, 8]. Trust primarily relates to medicine's contract with the public, i.e., that the public can trust the medical system to address their health and disease specific needs, provide inclusive and compassionate care, and be entrusted with the training of a competent workforce [9, 10]. It is the

extension of trust from the public, and from governments on behalf of the public, which enables medicine to exist as a self-regulating profession. A trustworthy assessment system is also important for physicians-in-training as it drives learning in medicine. An assessment system that builds and retains trust from the learners and the society is by definition accountable.

Decoupling assessment from its larger social context harms trust. The current state of assessment in medical education is characterized by a myriad of complex systems-level issues. Persistent group differences in standardized test performance and harmful bias in workplace-based assessment contribute to continued oppression of marginalized groups [11, 12]. This is due in large part to issues of socioeconomic disparities in resources and subsequent preparedness which are far more predominant in racially and ethnically marginalized student and trainee populations [12]. Issues of power, hierarchy and learner mistreatment additionally continue to exist in the clinical learning environment [13]. The current assessment system has contributed to the production of a workforce that is not representative of the population from the perspective of both demographics and lived experiences [14–17]. As a result, many graduates within our current training programs are not considered trustworthy to care for our family members for a host of reasons, yet they are permitted to enter unsupervised practice [18]. Lastly, health outcomes in many developed, resourced countries fail to justify the cost of healthcare, and demonstrate large disparities in which only certain sectors of the public experience positive outcomes [19]. These examples along the spectrum of education, training, and healthcare delivery provide ample justification for the current state of low trust in medicine as well as medical education assessment.

Through a lens of socio-constructivism and critical theory, considering both social context and power structures, we assert that the (re)establishment of trust should be foundational to the next era of assessment. In our actions and interventions, it is imperative that we address and build trust in assessment at a systems level – or our efforts to contextualize assessment, use data in innovative ways, and create systems of anti-oppression and justice will fail. In this paper, we first review the current state of assessment with attention paid to the social contextualization of assessment and its linkage to trust. We then discuss consequences that may come to pass should the current state of low trust carry over into the next era of assessment. We posit that trusting and trustworthy relationships can exist at individual as well as organizational and systems levels. Elements of individual and systems-level trust may differ, but core elements such as integrity, reliability and competence underlie both [20].

Finally, we propose a framework to build trust at multiple levels in a future assessment system, one that invites and supports professional and human growth and has the potential to position assessment as a fundamental component of renegotiating the social contract between medical education and the health of the public.

THE CURRENT STATE OF ASSESSMENT

As discussed previously, ideally our current assessment efforts should keep what is good from previous eras of assessment and discard or de-emphasize what is less valuable. We believe the current era of programmatic assessment (the best “picture” of performance is formed by considering multiple sources and types of data) and the assessment as judgment era (valuing qualitative assessment data and collective decision-making regarding whether performance is on track) that preceded it offer much good [21]. These approaches represent a move beyond a sole focus on the assessment of medical knowledge into the assessment of meaningful skills and behaviors and have the ability to capture competencies as they manifest in practice, in holistic or synthetic ways [22].

In our experiences, many programs are struggling to perfect assessment approaches introduced in prior assessment eras. For example, many programs are actively learning and sharing best practices for assimilating narrative data with ratings and optimizing clinical competency committee processes and structures [23–27]. However, they are doing this in an environment where national leaders and regulatory bodies continue to place emphasis on psychometrics introduced in the first era of assessment. This desire to tether the field to the thinking of past eras hides a reality that is laid bare for us – traditional ways of thinking about validity and making arguments for the veracity of our validity evidence that will be convincing for key stakeholders must change. This should lead us to more thoroughly consider argumentation in light of validity, which includes a greater emphasis on the consequences of our assessments [28].

Striving for equity and fairness in medical assessments demands a refined approach that emphasizes the optimization of narrative comments. However, equity and fairness in assessment is hindered by a narrow adherence to traditional validity concepts. Ensuring assessments effectively fit their intended purpose is key to both optimizing learning and making defensible decisions regarding progression [29]. Traditional psychometric approaches rely on standardized tests and quantitative measures to assess individuals’ traits, often emphasizing numerical scores. In contrast, newer psychometric approaches explore

qualitative contextual aspects of assessment, incorporating narrative data and dynamic techniques to provide a more comprehensive understanding of human behavior [30]. Evidence suggests that traditional psychometrics may inadvertently conceal biases, which is a significant concern in the context of medical evaluations. Additionally, narrative comments are increasingly cited as more useful than numerical ratings, as numerical performance is difficult to conceptualize or ascribe meaning, and these approaches often promote normative rankings and achievement orientation among learners [31–33]. The emphasis on these approaches promoting normative rankings and an achievement-oriented mindset among medical learners underscores the need to address these nuances. This contributes to a more comprehensive understanding of the challenges and potential improvements in achieving equity through medical assessments [34, 35].

Too often, assessment is reduced to a box-ticking exercise that offers little value for learners *by design* [36–38]. Learners are told that the assessment system is serving them well if they simply collect a specified number of assessments. Learners even talk about “needing to get 5 more EPAs done,” turning an approach designed for considering their work (EPAs) into a tool someone handed them that they have no clear use for [36, 38]. When learners view assessment in this manner, they tend to become frustrated with the assessment system within which they operate. If a program does not communicate the value of completing assessments, this then leads to their supervisors not understanding the value in assessments, and they as learners then do not understand the value in assessments. To all of these people working in the system, assessment is not about giving and receiving feedback for development, and it is certainly not about aiming that development at ensuring they are ready to meet the needs of patients. By implementing an assessment system based on a transformative idea like programmatic assessment in a superficial and transactional way, this has unintended negative consequences – this erodes trust in the system [39].

Exacerbating this situation, assessment efforts are too often a mechanism that can create distrust when assessment systems are not rigorously checked for fairness and equity and allow unwanted and harmful biases to run amuck [40]. When assessment does not strive to be fair and equitable, learners are harmed and trust is further eroded. This has become clear as biases in clerkship grades and test scores are identified, which create further issues within the leaky pipeline in medical education. As such, ensuring that assessment systems are designed with the intent to optimize equity while mitigating bias is key to ensuring a future that allows for the establishment of trust by all stakeholders in our assessments. The following sections analyze the barriers

to building trust in assessment and propose a way forward to rebuild this trust in the next era of assessment.

BARRIERS TO TRUST IN ASSESSMENT

What then is standing in our way when attempting to establish trust in assessment? Assessment systems wield power and some generate revenues (e.g., national board and licensing exams for which registration fees are charged). Acknowledging such drivers is critical to the conversation of trust. We can consider the system at micro (individual learner-assessor interactions), meso (an educational program), macro (an educational institution and affiliated health system environments), and global (regulatory bodies) levels.

At the micro level, the behaviors of individuals are driven by the incentives in and hierarchy of the larger system. Both learners and assessors are conscious of the implications of assessments. In many systems, learners incur significant financial and opportunity costs to pursue medical education. Given that small differences in assessment outcomes can have significant downstream impacts on opportunity and income [41], the assessment system can foster perverse incentives to hide developmental needs rather than foster growth and the development of competency. A consumeristic structure stokes learner frustration with seemingly haphazard clinical placements coupled with an apparent lack of support for assessors (insufficient resources, training and time for direct observation and feedback) [42]. Sensitive to an apparent lack of construct validity [43, 44], learners may feel driven to “gaming” the system and assessors may empathize [45].

At the meso level, educational programs are under performance pressures. Evaluation and ranking of educational programs – often via metrics that lack alignment with educational outcomes (e.g. US News and World Report rankings of medical schools) – impact funds flow and talent recruitment [46, 47]. Programs are often under-resourced, leading to structures built around convenience of managing cohorts rather than tailoring to needs of individuals. This can promote an inclination among programs to defend against potential “weak” performers who may require more support. Additionally, litigation challenges to negative assessment decisions, such as probation or dismissal, create a disincentive to address competency concerns. Lacking structures and support for remediation, programs may develop an unspoken reliance on future elements or phases of the assessment system (e.g. certification examinations instead of rigorous programmatic assessment) to address serious issues, undermining the process of a self-regulating profession [48].

At the macro level, there is frequently a lack of alignment of incentives along the training program continuum (undergraduate medical education (UME), graduate medical education (GME) and continuing professional development (CPD)) and the health systems in which clinical training occurs. For example, health systems driven by throughput and margins may develop a short-term focus on efficiency at the expense of the long term benefits of the education mission. It is incumbent upon the medical education community to describe the value proposition of strengthening the integration of learning and work such that it is clear how these integrate and complement one another. Communicating the expected benefits to patient outcomes and clinician well-being that could be realized by deliberately developmental organizations may facilitate alignment [49].

At the global level, accreditors and licensing bodies developed under a historic deficit orientation to assessment and assumed an exclusionary stance. This means that accreditors and licensing bodies focused on ensuring that they theoretically excluded those who are not competent to care for patients from becoming board certified and/or licensed. Accreditors remain heavily focused on assessing consistency of process, despite the premise of competency-based medical education (CBME) demonstrating that process should vary to support desired outcomes. A lack of a shared mental model of competency across the continuum makes it challenging for licensing bodies to support local solutions while simultaneously fulfilling their function of external audit and accountability. Although steps have been taken to redesign processes to align ongoing certification with desired educational and performance outcomes, licensing and accreditation remain perceived by many as burdensome, costly, and often irrelevant to real world needs [50–52].

RESTORING TRUST IN ASSESSMENT

As we hope to move medical education into a new era of assessment, building a trustworthy system is imperative to success. Given that assessment data may be used across the continuum of medical education, we need to focus on building trust throughout all levels of this system [53–56]. We propose an assessment paradigm that is focused on trust at four levels: trust in the individual providing the assessment data (microsystem), trust in program or school leadership (mesosystem), trust in the institution and health system (macrosystem), and trust in the accrediting and licensing bodies (global assessment system). Using the framework developed by Robert Hurley to design trustworthy organizations, Table 1 provides several examples of each [20].

	MICRO (INDIVIDUAL)	MESO (PROGRAM)	MACRO (SYSTEM)	GLOBAL (ACCREDITING AND LICENSING BODIES)
Common Values	<ul style="list-style-type: none"> - Supervisor and learner dyad both value patient autonomy, are committed to patient safety, and thus prioritize shared decision making in clinical encounters 	<ul style="list-style-type: none"> - Program is transparent in selection, advancement, and graduation processes; shares their mission; works to ensure equity in assessment and selection and is transparent in these efforts, and is committed to following these principles 	<ul style="list-style-type: none"> - Learners aligned with institution's mission and institution consistently delivers on its commitments to learners 	<ul style="list-style-type: none"> - Learner and certification and licensing bodies have a shared mental model of outcomes that define a competent physician
Aligned Interests	<ul style="list-style-type: none"> - Supervisor and learner dyad share interest in learner growth and improvement - Learners and supervisors coproduce assessments and subsequent feedback discussions and learning opportunities to target individual needs 	<ul style="list-style-type: none"> - Program seeks feedback from learners, makes changes with learners' best interests in mind - Workflows and supervisor financial incentives support effort devoted to assessment and associated feedback 	<ul style="list-style-type: none"> - Institution is clear in their commitment to education, balancing service demands on learners by employing advanced practice providers and other staff - Strategies in place to minimize tuition and learner debt 	<ul style="list-style-type: none"> - Certification and licensing bodies and learners share interest in ensuring the provision of safe, effective, and high-quality patient care and share a vision of the relevant patient outcomes across the educational continuum - Developmental organizations recognize the business case for financial investment in education at all levels of continuum
Benevolence	<ul style="list-style-type: none"> - Supervisor frames assessment and the feedback that follows in context of a growth mindset, does not dwell unduly on negative outcomes 	<ul style="list-style-type: none"> - Program community supports learners who are going through challenges 	<ul style="list-style-type: none"> - Institutions treat learners as colleagues, including them in employee appreciation events, bonus distribution, decision-making, etc. 	<ul style="list-style-type: none"> - Certification and licensing bodies make accommodations when learners have extenuating circumstances (e.g., appropriately consider test accommodations, extensions, and deferments) and also ensure accommodations do not need to be requested if they are consistent with common situations (e.g., time and space for pumping for breastfeeding mothers)
Competence / Capability	<ul style="list-style-type: none"> - Learner and supervisors have a shared mental model of competence, resulting in learners respecting supervisors' clinical knowledge, experience, and observations which lead to the assessment and subsequent feedback - Supervisors normalize the developmental process and make sure learners know it is acceptable to not know something yet, sharing their previous and ongoing learning efforts whenever possible 	<ul style="list-style-type: none"> - Learners expect program leadership to make necessary changes with their best interests in mind - Program provides opportunity for graduated responsibility and progressive autonomy 	<ul style="list-style-type: none"> - GME and hospital leadership conduct themselves in a manner that allows learners to have confidence in them - Institutional leaders use learner centered data as part of continuous quality improvement (CQI) process 	<ul style="list-style-type: none"> - Licensing and certifying bodies have activities and processes that accurately assess learners' knowledge base and relevant outcomes - Certifying and licensing assessments are designed and rigorously tested in a multitude of contexts and with a diverse group of learners to minimize bias

(Contd.)

	MICRO (INDIVIDUAL)	MESO (PROGRAM)	MACRO (SYSTEM)	GLOBAL (ACCREDITING AND LICENSING BODIES)
Predictability & Integrity	<ul style="list-style-type: none"> - Supervisor provides necessary individualized support to all learners 	<ul style="list-style-type: none"> - Program handles resident matters with confidentiality and discretion - Concerns raised by leadership or learner are handled promptly and professionally - Programmatic assessment is implemented equitably across all learners 	<ul style="list-style-type: none"> - Institution or GME office is consistent with treatment of learners across training programs (e.g., duty hour violations or professionalism lapses are handled consistently for surgery and medical residents) - Outcomes for all learners are clearly defined and transparent in how they are assessed 	<ul style="list-style-type: none"> - Regulatory policies are clearly written and openly shared on website so that assessed outcomes and standards are transparent and accountable to learners and the public
Communication	<ul style="list-style-type: none"> - There is open dialogue between learner and supervisor within feedback conversations - Supervisor maintains and encourages open communication with the learner through multiple avenues (page, email, phone number) - Supervisor is trained in providing feedback and engages in reflective practice to identify and mitigate own blind spots 	<ul style="list-style-type: none"> - Program leadership provides regular forum to communicate with residents outside of email and phone (e.g., debriefs, feedback sessions at retreats, anonymous concern reporting) 	<ul style="list-style-type: none"> - Institutional leaders are accessible to learners (e.g., they hold listening sessions or open houses to get to know learners) 	<ul style="list-style-type: none"> - Learners have a seat at the table with licensing and regulatory bodies and have an avenue to provide frequent and rigorous feedback to facilitate continuous quality improvement of assessments that is incorporated into regulatory process

Table 1 Examples of trust-promoting interactions between learners and the medical education system across levels.

To enhance trust, we posit that the context of assessment in the next era must be deeply and skillfully embedded in relationships. We assert that the most immediately important relationship is the one between the learner and the individual providing them with feedback after an assessment (assessor). While the assessment source may be a teacher (faculty member), they may also be a peer or team member (fellow resident, nurse, patient's family member). It is imperative that any assessment comes from a valid source; i.e., that the person providing the assessment be trusted by the learner and be considered credible [13, 57]. However, the learner's dependent position in the medical education hierarchy often complicates assessment from supervisory sources. Previous research has shown that learner perception of how supervisors used their power was a key influence on learner trust in their supervisors and subsequent engagement in assessment for learning [13, 58].

Additionally, integrating reflection into supervisor behavior enhances learner trust and participation in their own learning. For example, the R2C2 Feedback Model promotes feedback after an assessment as an interaction in which learner engagement, supportive relationships, reflection, and cooperative planning are emphasized [59–61]. One challenge that exists in large institutions, particularly in the era of remote work where supervisors may leave campus after rounds to work on academic pursuits, is that frequent communication must now be scheduled rather than happening organically which under significant time pressures can hinder formation of close bonds between learner and teacher. Furthermore, learners are often pulled between multiple different learning environments which can make fully engaging in each clinical space difficult. Ideally, learners – with the help of a coach – would organize assessment data to reach specific learning goals (i.e., for a resident

interested in a procedural-based fellowship, the coach may help them focus heavily on data related to procedural competence). Learners should also be encouraged to co-produce assessment data that is important to them [62]. Based on identified goals and values, learners should seek and document self-assessments and assessments from colleagues. This model will also aid in lowering the perceived stakes of frequent assessments and focus the learner's attention on growth and development.

Sharing common values and aligned interests, displaying clinical competence, and having good communication skills all enhance a teacher's ability to be viewed as valid and credible by the learner (Table 1). Properly weighing different sources of assessment FOR learning when using individual data points for aggregate assessment OF learning is important [63]. For example, a supervising physician whose assessment of a learner is based solely on one negative patient interaction that they witnessed is less likely to be viewed as trustworthy than one which was compiled from numerous direct observations over time. Transparency and consistency across assessment mediums is also critical; that is verbal feedback provided after an assessment should be consistent with written feedback to achieve maximum trust. When delivering verbal feedback and corresponding written narratives, we might recommend employing the four tenets of the R2C2 approach as a model for providing facilitated feedback after an assessment: 1. Building rapport and relationship; 2. Exploring reactions to and perceptions of the assessment; 3. Explore understanding of the content of the assessment; and 4. Coaching for performance change [59]. While supervising physicians often do not receive any formal training in providing feedback to residents, even a short supervisor training session can significantly improve the quality of their feedback [64].

As we move through the assessment system, these same factors are important at the program (mesosystem) level but may be harder to attain as the learner must now determine the program's trustworthiness based on the collective values and competence of a larger number of faculty members. A program that is able to successfully convey its investment in the development and wellbeing of its learners can cultivate a psychologically safe learning environment in which learners can focus on growth and development. Including assessment in a program's core values and mission statements can help clearly align interests of the learners and the program. For example, indicating that a program strives to graduate board certified physicians indicates the value that a given program places on certification exams as an assessment tool.

The program must then reliably deliver on its commitments and continuously assess its success in doing so to continue to foster trust with learners. Once a learner has established their goals within the training program and program leadership

has reviewed and approved this, there must be a mutual investment in achieving those goals meaning the program and supervisors must partner with and support the learner in achieving those goals via individualized learning approaches. In this relationship, the learner is responsible for maintaining a growth mindset in which they see shortcomings as an opportunity to learn and improve [65–67]. Programs must provide the learner with ample opportunities for learning and must ensure equity and fairness in assessments and promotion decisions by working to mitigate bias and iteratively auditing assessments for potential biases.

At the macro level, institutions must commit to holding themselves, individual programs, and the health system which they oversee accountable to ensuring the attainment and maintenance of a trustworthy assessment system. This is imperative to the success of the overall functioning of the institution as well as to ensuring the optimal coexistence of learning and safe and effective patient care. This first requires that institutions, organizations, and health systems view themselves as key stakeholders in the educational system and commit to participating in the provision of education. Transparent coproduction of expectations between programs and institutions to build trust and alignment around educational outcomes particularly tied to learning and patient care instead of profits and market trends is a key step in that direction. Additionally, committing to and promoting psychological safety and equity in assessment at a systems level within the learning environment shows investment in common values between the system, programs, and learners [68, 69]. Additionally, medical educators overseeing assessment programs should monitor trust across stakeholders and contribute to a broader research agenda to foster trust across institutions.

The training of learners who can safely and effectively care for patients is the primary aim of assessment at the macro level. As such, consistent, transparent regulatory support, check-ins, and communication that involves all stakeholders including programs and learners is key to demonstrate a commitment to that primary aim. This requires consistent quality assurance of the educational system, a visible commitment to equity and fairness in assessment, and recurrent discussions of the support structures and coaching available to help learners through their training and assessment in order to help them meet the clear and agreed upon outcomes required for progression. A system that consistently delivers on competence across all programs and demonstrates a commitment to both learners and patients is one that will garner the trust of all stakeholders under its umbrella.

Finally, at the global level, the power that regulatory bodies hold has historically led to inherent distrust by learners and program leadership [70–74]. The lack of direct

contact with these assessors inhibits the ability to build trust through communication and personal relationships. Regulators at the global level must build trust through a dedication to integrity, equity, fairness, and predictability. However, regulatory and licensing bodies often view this process through a different lens – that of ensuring patient and societal safety [75, 76]. Flipping this concept of trust at the global level to their viewpoint, we also owe it to another stakeholder, our patients, to ensure that society trusts the healthcare system’s ability to self-regulate and ensure that they are cared for by competent medical professionals. Society ultimately determines how trustworthy our assessments are by how much trust they have in the healthcare system.

To that end, we must reconsider our approach to licensing and certification through both of these lenses. This starts by regulatory bodies sharing priorities in assessment with those of learners and practicing clinicians as well as programs. Outcomes for certification must center around patients and society but also be relevant to learners and programs. Alignment between learners, programs, and regulatory bodies can be achieved by creating a transparent shared vision and understanding around the purpose of each assessment and its relationship to relevant outcomes to all stakeholders, particularly learners and patients. A frequent back and forth with programs and learners that builds a culture of mutual trust and respect will mitigate the feeling that certification and licensing is a “black box” that learners cannot trust. This will additionally require ensuring equity in assessment, which will require ensuring access to learning resources, mitigation of bias in assessment, and rigorous audits of systems to optimize equity between a variety of groups of learners at all levels. There must be meaningful and rigorous outcomes; however all learners must be supported in meeting those outcomes. Graded certification with these outcomes in mind will aid in trust, particularly with learners, while also ensuring patients’ needs are met [77]. Finally, transparent continuous quality improvement must be employed by regulatory and licensing bodies to demonstrate accountability to learners, programs, and society. This can be achieved by the coproduction and sharing of outcomes and measures by which regulatory and licensing bodies assess programs and learners, such that they are prepared to train future healthcare providers who are competent to care for patients and society respectively.

CONCLUSIONS

The success of the next era of assessment hinges upon mutual trust from and between all levels: learners, program leaders, healthcare systems, accreditation bodies, and


patients. It is imperative that we learn from current barriers to trust to develop trustworthy assessment systems that focus on common values and aligned interests, clinical competence, and honest communication. Future work to develop new assessment schema should incorporate coproduction across all levels of the system to ensure equity and fairness for learners while maintaining focus on our profession’s ultimate responsibility to provide high-quality care to the patients we serve.

COMPETING INTERESTS

Dr. Lomis is an employee of the American Medical Association. The opinions expressed in this article are those of the authors and do not necessarily reflect American Medical Association policy.

Dr. Barone is an employee of NBME. The opinions expressed in this article are those of the authors and do not necessarily reflect NBME policy.

AUTHOR AFFILIATIONS

Holly A. Caretta-Weyer, MD, MHPE  orcid.org/0000-0002-9783-5797

Clinical Associate Professor and Associate Residency Director in the Department of Emergency Medicine, Stanford University School of Medicine, Palo Alto, California, USA

Alina Smirnova, MD, PhD  orcid.org/0000-0003-4491-3007

Clinical assistant professor, Department of Family Medicine, University of Calgary, Calgary, Alberta, Canada; Adjunct assistant professor, Kern Institute for the Transformation of Medical Education, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

Michael A. Barone, MD  orcid.org/0000-0002-4724-784X

Vice president, competency-based assessment, NBME, Philadelphia, Pennsylvania, USA; Adjunct associate professor of pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Jason R. Frank, MD  orcid.org/0000-0002-6076-0146

Professor of emergency medicine, Department of Emergency Medicine, University of Ottawa, Ottawa, Ontario, CA

Tina Hernandez-Boussard, PhD  orcid.org/0000-0002-7937-9564

Professor of medicine (biomedical informatic), of biomedical data, and of surgery, Stanford University School of Medicine, Palo Alto, California, USA

Dana Levinson, MPH

Chief Program Officer, the Josiah Macy Jr Foundation, Philadelphia, Pennsylvania, USA

Kiki M. J. M. H. Lombarts, PhD  orcid.org/0000-0001-6167-0620

Professor of professional performance, Department of Medical Psychology, Amsterdam University Medical Centers, University of Amsterdam, NL; Amsterdam Public Health research institute,

Amsterdam, NL

Kimberly D. Lomis, MD  orcid.org/0000-0002-3504-6776

Vice president, Undergraduate Medical Education Innovations, American Medical Association, Chicago, Illinois, USA

Abigail Martini, MSMed  orcid.org/0000-0003-1860-8113

Clinical research coordinator with emergency medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

Daniel J. Schumacher, MD, PhD  orcid.org/0000-0001-5507-8452

Professor of pediatrics, Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center/University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

David A. Turner, MD

Vice President for Competency-Based Medical Education (CBME) at the American Board of Pediatrics, Chapel Hill, North Carolina, USA

Abigail Schuh, MD  orcid.org/0000-0002-6422-2361

Associate professor of pediatrics, Division of Emergency Medicine, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

REFERENCES

1. **Sklar DP, McMahon GT.** Trust Between Teachers and Learners. *JAMA*. 2019; 321(22): 2157. DOI: <https://doi.org/10.1001/jama.2018.22130>
2. **Helmreich RL.** On error management: lessons from aviation. *BMJ*. 2000; 320(7237): 781–785. DOI: <https://doi.org/10.1136/bmj.320.7237.781>
3. **Rigamonti D, Rigamonti KH.** Achieving and Maintaining Safety in Healthcare Requires Unwavering Institutional and Individual Commitments. *Cureus*. Published online February 7, 2021. DOI: <https://doi.org/10.7759/cureus.13192>
4. **Hafferty FW, O'Brien BC, Tilburt JC.** Beyond High-Stakes Testing: Learner Trust, Educational Commodification, and the Loss of Medical School Professionalism. *Acad Med*. 2020; 95(6): 833–837. DOI: <https://doi.org/10.1097/ACM.00000000000003193>
5. **Moynahan KF.** The Current Use of United States Medical Licensing Examination Step 1 Scores: Holistic Admissions and Student Well-Being Are in the Balance. *Acad Med*. 2018; 93(7): 963–965. DOI: <https://doi.org/10.1097/ACM.00000000000002101>
6. **Schuwirth LWT, Van Der Vleuten CPM.** A history of assessment in medical education. *Adv Health Sci Educ*. 2020; 25(5): 1045–1056. DOI: <https://doi.org/10.1007/s10459-020-10003-0>
7. **Geraghty JR, Ocampo RG, Liang S, et al.** Medical Students' Views on Implementing the Core EPAs: Recommendations From Student Leaders at the Core EPAs Pilot Institutions. *Acad Med*. 2021; 96(2): 193–198. DOI: <https://doi.org/10.1097/ACM.00000000000003793>
8. **Yengo-Kahn AM, Baker CE, Lomis AKD.** Medical Students' Perspectives on Implementing Curriculum Change at One Institution. *Acad Med*. 2017; 92(4): 455–461. DOI: <https://doi.org/10.1097/ACM.0000000000001569>
9. **Phillips RL, George BC, Holmboe ES, Bazemore AW, Westfall JM, Bitton A.** Measuring Graduate Medical Education Outcomes to Honor the Social Contract. *Acad Med J Assoc Am Med Coll*. 2022; 97(5): 643–648. DOI: <https://doi.org/10.1097/ACM.00000000000004592>
10. **Cruess RL, Cruess SR.** Professionalism, Communities of Practice, and Medicine's Social Contract. *J Am Board Fam Med*. 2020; 33(Supplement): S50–S56. DOI: <https://doi.org/10.3122/jabfm.2020.S1.190417>
11. **Rubright JD, Jodoin M, Barone MA.** Examining Demographics, Prior Academic Performance, and United States Medical Licensing Examination Scores. *Acad Med*. 2019; 94(3): 364–370. DOI: <https://doi.org/10.1097/ACM.00000000000002366>
12. **Boatright D, Anderson N, Kim JG, et al.** Racial and Ethnic Differences in Internal Medicine Residency Assessments. *JAMA Netw Open*. 2022; 5(12): e2247649. DOI: <https://doi.org/10.1001/jamanetworkopen.2022.47649>
13. **Castanelli DJ, Weller JM, Molloy E, Bearman M.** Trust, power and learning in workplace-based assessment: The trainee perspective. *Med Educ*. 2022; 56(3): 280–291. DOI: <https://doi.org/10.1111/medu.14631>
14. **Moy E, Bartman BA.** Physician race and care of minority and medically indigent patients. *JAMA*. 1995; 273(19): 1515–1520. DOI: <https://doi.org/10.1001/jama.1995.03520430051038>
15. **Laveist TA, Nuru-Jeter A.** Is doctor-patient race concordance associated with greater satisfaction with care? *J Health Soc Behav*. 2002; 43(3): 296–306. DOI: <https://doi.org/10.2307/3090205>
16. **LaVeist TA, Nuru-Jeter A, Jones KE.** The association of doctor-patient race concordance with health services utilization. *J Public Health Policy*. 2003; 24(3–4): 312–323. DOI: <https://doi.org/10.2307/3343378>
17. **Cohen JJ, Gabriel BA, Terrell C.** The case for diversity in the health care workforce. *Health Aff Proj Hope*. 2002; 21(5): 90–102. DOI: <https://doi.org/10.1377/hlthaff.21.5.90>
18. **Jonker G, Ochtman A, Marty AP, Kalkman CJ, Ten Cate O, Hoff RG.** Would you trust your loved ones to this trainee? Certification decisions in postgraduate anaesthesia training. *Br J Anaesth*. 2020; 125(5): e408–e410. DOI: <https://doi.org/10.1016/j.bja.2020.07.009>
19. **Schneider EC, Shah A, Doty MM, Tikkanen R, Fields K, Williams RD, III.** *Mirror, Mirror 2021: Reflecting Poorly. Health Care in the US Compared to Other High-Income Countries*. Commonwealth Fund; 2021. Accessed July 16, 2022. <https://www.commonwealthfund.org/publications/fund-reports/2021/aug/mirror-mirror-2021-reflecting-poorly>.

20. **Hurley RF, Gillespie N, Ferrin DL, Dietz G.** Designing Trustworthy Organizations. *MIT Sloan Manag Rev.* 54(4): 74–82.
21. **Vleuten C, Lindemann I, Schmidt L.** Programmatic assessment: the process, rationale and evidence for modern evaluation approaches in medical education. *Med J Aust.* 2018; 209(9): 386–388. DOI: <https://doi.org/10.5694/mja17.00926>
22. **Ten Cate O, Carraccio C, Damodaran A,** et al. Entrustment Decision Making: Extending Miller's Pyramid. *Acad Med.* 2021; 96(2): 199–204. DOI: <https://doi.org/10.1097/ACM.0000000000003800>
23. **Ginsburg S, Van Der Vleuten CPM, Eva KW.** The Hidden Value of Narrative Comments for Assessment: A Quantitative Reliability Analysis of Qualitative Data. *Acad Med.* 2017; 92(11): 1617–1621. DOI: <https://doi.org/10.1097/ACM.0000000000001669>
24. **Ginsburg S, Regehr G, Lingard L, Eva KW.** Reading between the lines: faculty interpretations of narrative evaluation comments. *Med Educ.* 2015; 49(3): 296–306. DOI: <https://doi.org/10.1111/medu.12637>
25. **Edgar L, Jones MD, Harsy B, Passiment M, Hauer KE.** Better Decision-Making: Shared Mental Models and the Clinical Competency Committee. *J Grad Med Educ.* 2021; 13(2s): 51–58. DOI: <https://doi.org/10.4300/JGME-D-20-00850.1>
26. **Hauer KE, Cate OT, Boscardin CK,** et al. Ensuring Resident Competence: A Narrative Review of the Literature on Group Decision Making to Inform the Work of Clinical Competency Committees. *J Grad Med Educ.* 2016; 8(2): 156–164. DOI: <https://doi.org/10.4300/JGME-D-15-00144.1>
27. **Kinnear B, Warm EJ, Hauer KE.** Twelve tips to maximize the value of a clinical competency committee in postgraduate medical education. *Med Teach.* 2018; 40(11): 1110–1115. DOI: <https://doi.org/10.1080/0142159X.2018.1474191>
28. **Kinnear B, Schumacher DJ, Driessen EW, Varpio L.** How argumentation theory can inform assessment validity: A critical review. *Med Educ.* 2022; 56(11): 1064–1075. DOI: <https://doi.org/10.1111/medu.14882>
29. **Ross S, Hauer KE, Wycliffe-Jones K,** et al. Key considerations in planning and designing programmatic assessment in competency-based medical education. *Med Teach.* 2021; 43(7): 758–764. DOI: <https://doi.org/10.1080/0142159X.2021.1925099>
30. **Lomis KD, Russell RG, Davidson MA,** et al. Competency milestones for medical students: Design, implementation, and analysis at one medical school. *Med Teach.* 2017; 39(5): 494–504. DOI: <https://doi.org/10.1080/0142159X.2017.1299924>
31. **Hodges B.** Assessment in the post-psychometric era: learning to love the subjective and collective. *Med Teach.* 2013; 35(7): 564–568. DOI: <https://doi.org/10.3109/0142159X.2013.789134>
32. **Ten Cate O, Regehr G.** The Power of Subjectivity in the Assessment of Medical Trainees. *Acad Med J Assoc Am Med Coll.* 2019; 94(3): 333–337. DOI: <https://doi.org/10.1097/ACM.0000000000002495>
33. **Schumacher DJ, Michelson C, Poynter S,** et al. Thresholds and interpretations: How clinical competency committees identify pediatric residents with performance concerns. *Med Teach.* 2018; 40(1): 70–79. DOI: <https://doi.org/10.1080/0142159X.2017.1394576>
34. **Norman GR, Vleuten CPM, Graaff E.** Pitfalls in the pursuit of objectivity: issues of validity, efficiency and acceptability. *Med Educ.* 1991; 25(2): 119–126. DOI: <https://doi.org/10.1111/j.1365-2923.1991.tb00037.x>
35. **Vleuten CPM, Norman GR, Graaff E.** Pitfalls in the pursuit of objectivity: issues of reliability. *Med Educ.* 1991; 25(2): 110–118. DOI: <https://doi.org/10.1111/j.1365-2923.1991.tb00036.x>
36. **Ott MC, Pack R, Cristancho S, Chin M, Van Koughnett JA, Ott M.** “The Most Crushing Thing”: Understanding Resident Assessment Burden in a Competency-Based Curriculum. *J Grad Med Educ.* 2022; 14(5): 583–592. DOI: <https://doi.org/10.4300/JGME-D-22-00050.1>
37. **Sehlbach C, Govaerts MJB, Mitchell S, Rohde GGU, Smeenk FWJM, Driessen EW.** Box-ticking and Olympic high jumping – Physicians' perceptions and acceptance of national physician validation systems. *Med Teach.* 2018; 40(9): 886–891. DOI: <https://doi.org/10.1080/0142159X.2018.1470320>
38. **Szulewski A, Braund H, Dagnone DJ,** et al. The Assessment Burden in Competency-Based Medical Education: How Programs Are Adapting. *Acad Med.* 2023; 98(11): 1261–1267. DOI: <https://doi.org/10.1097/ACM.0000000000005305>
39. **Lomis KD, Mejicano GC, Caverzagie KJ, Monrad SU, Pusic M, Hauer KE.** The critical role of infrastructure and organizational culture in implementing competency-based education and individualized pathways in undergraduate medical education. *Med Teach.* 2021; 43(sup2): S7–S16. DOI: <https://doi.org/10.1080/0142159X.2021.1924364>
40. **Hauer KE, Park YS, Bullock JL, Tekian A.** “My Assessments Are Biased!” Measurement and Sociocultural Approaches to Achieve Fairness in Assessment in Medical Education. *Acad Med.* 2023; 98(8S): S16–S27. DOI: <https://doi.org/10.1097/ACM.0000000000005245>
41. **Teherani A, Hauer KE, Fernandez A, King TE, Lucey C.** How Small Differences in Assessed Clinical Performance Amplify to Large Differences in Grades and Awards: A Cascade With Serious Consequences for Students Underrepresented in Medicine. *Acad Med.* 2018; 93(9): 1286–1292. DOI: <https://doi.org/10.1097/ACM.0000000000002323>
42. **Beck Dallaghan GL, Lin X, Melvin JK, Golding J, Steiner B, Kulkarni V.** Maximizing clinical rotation placements for US medical students: exploring an optimization model. *Med Educ Online.* 2022; 27(1): 2024488. DOI: <https://doi.org/10.1080/10872981.2021.2024488>

43. **Messick S.** Validity. In: *Educational Measurement*. 3rd ed. American Council on Education: Macmillan Pub. Co.; Collier Macmillan Publishers; 1989: 13–103.
44. **Cronbach LJ, Meehl PE.** Construct validity in psychological tests. *Psychol Bull.* 1955; 52(4): 281–302. DOI: <https://doi.org/10.1037/h0040957>
45. **Armstrong EG, Mackey M, Spear SJ.** Medical Education as a Process Management Problem: *Acad Med.* 2004; 79(8): 721–728. DOI: <https://doi.org/10.1097/00001888-200408000-00002>
46. **Conference Participants.** Josiah Macy Jr. Foundation Conference on COVID-19 and the Impact on Medical and Nursing Education: Conference Recommendations Report. *Acad Med.* 2022; 97(3S): S3–S11. DOI: <https://doi.org/10.1097/ACM.0000000000004506>
47. **Jaschik S.** Harvard Med School Drops ‘U.S. News.’ Inside Higher Ed. Published January 22, 2023. Accessed November 1, 2023. <https://www.insidehighered.com/admissions/article/2023/01/23/harvard-med-school-drops-us-news-will-others-follow>.
48. **Dudek NL, Marks MB, Regehr G.** Failure to Fail: The Perspectives of Clinical Supervisors. *Acad Med.* 2005; 80(Supplement): S84–S87. DOI: <https://doi.org/10.1097/00001888-200510001-00023>
49. **Thoma B, Caretta-Weyer H, Schumacher DJ, et al.** Becoming a deliberately developmental organization: Using competency based assessment data for organizational development. *Med Teach.* 2021; 43(7): 801–809. DOI: <https://doi.org/10.1080/0142159X.2021.1925100>
50. **University of Colorado Anschutz School of Medicine, Price DW, Davis DA, The AXDEV Group, Filerman GL, Atlas Health Foundation.** “Systems-Integrated CME”: The Implementation and Outcomes Imperative for Continuing Medical Education in the Learning Health Care Enterprise. *NAM Perspect.* 2021; 11(10). DOI: <https://doi.org/10.31478/202110a>
51. **Price DW, Biernacki H, Nora LM.** Can Maintenance of Certification Work? Associations of MOC and Improvements in Physicians’ Knowledge and Practice. *Acad Med.* 2018; 93(12): 1872–1881. DOI: <https://doi.org/10.1097/ACM.0000000000002338>
52. **McMahon GT.** Inspiring Curiosity and Restoring Humility: The Evolution of Competency-Based Continuing Medical Education. *Acad Med.* 2018; 93(12): 1757–1759. DOI: <https://doi.org/10.1097/ACM.0000000000002339>
53. **Wilhite JA, Altschuler L, Zabar S, Gillespie C, Kalet A.** Development and maintenance of a medical education research registry. *BMC Med Educ.* 2020; 20(1): 199. DOI: <https://doi.org/10.1186/s12909-020-02113-5>
54. **Wancata LM, Morgan H, Sandhu G, Santen S, Hughes DT.** Using the ACMGE Milestones as a Handover Tool From Medical School to Surgery Residency. *J Surg Educ.* 2017; 74(3): 519–529. DOI: <https://doi.org/10.1016/j.surg.2016.10.016>
55. **Schiller JH, Burrows HL, Fleming AE, Keeley MG, Wozniak L, Santen SA.** Responsible Milestone-Based Educational Handover With Individualized Learning Plan From Undergraduate to Graduate Pediatric Medical Education. *Acad Pediatr.* 2018; 18(2): 231–233. DOI: <https://doi.org/10.1016/j.acap.2017.09.010>
56. **Sozener CB, Lypson ML, House JB, et al.** Reporting Achievement of Medical Student Milestones to Residency Program Directors: An Educational Handover. *Acad Med.* 2016; 91(5): 676–684. DOI: <https://doi.org/10.1097/ACM.0000000000000953>
57. **Abruzzo D, Sklar DP, McMahon GT.** Improving Trust Between Learners and Teachers in Medicine: *Acad Med.* 2019; 94(2): 147–150. DOI: <https://doi.org/10.1097/ACM.0000000000002514>
58. **O’Connor A, McCurtin A, Cantillon P, McGarr O.** Illusions of specificity in power-laden clinical performance assessment. *Med Teach.* 2018; 40(3): 313–314. DOI: <https://doi.org/10.1080/0142159X.2017.1340637>
59. **Sargeant J, Lockyer J, Mann K, et al.** Facilitated Reflective Performance Feedback: Developing an Evidence- and Theory-Based Model That Builds Relationship, Explores Reactions and Content, and Coaches for Performance Change (R2C2). *Acad Med.* 2015; 90(12): 1698–1706. DOI: <https://doi.org/10.1097/ACM.0000000000000809>
60. **Lockyer J, Armson H, Könings KD, et al.** In-the-Moment Feedback and Coaching: Improving R2C2 for a New Context. *J Grad Med Educ.* 2020; 12(1): 27–35. DOI: <https://doi.org/10.4300/JGME-D-19-00508.1>
61. **Armson H, Lockyer JM, Zetkovic M, Könings KD, Sargeant J.** Identifying coaching skills to improve feedback use in postgraduate medical education. *Med Educ.* 2019; 53(5): 477–493. DOI: <https://doi.org/10.1111/medu.13818>
62. **Englander R, Holmboe E, Batalden P, et al.** Coproducing Health Professions Education: A Prerequisite to Coproducing Health Care Services? *Acad Med.* 2020; 95(7): 1006–1013. DOI: <https://doi.org/10.1097/ACM.0000000000003137>
63. **Kinnear B, Warm EJ, Caretta-Weyer H, et al.** Entrustment Unpacked: Aligning Purposes, Stakes, and Processes to Enhance Learner Assessment. *Acad Med.* 2021; 96(7S): S56–S63. DOI: <https://doi.org/10.1097/ACM.0000000000004108>
64. **Renting N, Jaarsma D, Borleffs JC, Slaets JPJ, Cohen-Schotanus J, Gans ROB.** Effectiveness of a supervisor training on quality of feedback to internal medicine residents: a controlled longitudinal multicentre study. *BMJ Open.* 2023; 13(9): e076946. DOI: <https://doi.org/10.1136/bmjopen-2023-076946>
65. **Dweck C.** *Mindset: The New Psychology of Success.* Ballantine Books; 2016.

66. **Richardson D, Kinnear B, Hauer KE**, et al. Growth mindset in competency-based medical education. *Med Teach*. 2021; 43(7): 751–757. DOI: <https://doi.org/10.1080/0142159X.2021.1928036>
67. **Han SJ, Stieha V**. Growth Mindset for Human Resource Development: A Scoping Review of the Literature with Recommended Interventions. *Hum Resour Dev Rev*. 2020; 19(3): 309–331. DOI: <https://doi.org/10.1177/1534484320939739>
68. **Carmeli A, Gittell JH**. High-quality relationships, psychological safety, and learning from failures in work organizations. *J Organ Behav*. 2009; 30(6): 709–729. DOI: <https://doi.org/10.1002/job.565>
69. **Clark TR**. *The 4 Stages of Psychological Safety: Defining the Path to Inclusion and Innovation*. First edition. Berrett-Koehler Publishers, Inc; 2020.
70. **Buscemi D, Wang H, Phy M, Nugent K**. Maintenance of certification in Internal Medicine: participation rates and patient outcomes. *J Community Hosp Intern Med Perspect*. 2013; 2(4): 19753. DOI: <https://doi.org/10.3402/jchimp.v2i4.19753>
71. **Sandhu AT, Dudley RA, Kazi DS**. A Cost Analysis of the American Board of Internal Medicine's Maintenance-of-Certification Program. *Ann Intern Med*. 2015; 163(6): 401–408. DOI: <https://doi.org/10.7326/M15-1011>
72. **Gesundheit N**. A Crisis of Trust Between U.S. Medical Education and the National Board of Medical Examiners. *Acad Med*. 2020; 95(9): 1300–1304. DOI: <https://doi.org/10.1097/ACM.0000000000003131>
73. **Carmody JB, Rajasekaran SK**. On Step 1 Mania, USMLE Score Reporting, and Financial Conflict of Interest at the National Board of Medical Examiners. *Acad Med*. 2020; 95(9): 1332–1337. DOI: <https://doi.org/10.1097/ACM.0000000000003126>
74. **Martin K**. ABIM needs to look in the mirror when discussing lack of trust. *Medical Economics*. Published May 10, 2017. Accessed November 9, 2023. <https://www.medicaleconomics.com/view/abim-needs-look-mirror-when-discussing-lack-trust>.
75. **Professionalism and Medicine's Social Contract with Society**. *AMA J Ethics*. 2004; 6(4). DOI: <https://doi.org/10.1001/virtualmentor.2004.6.4.msoc1-0404>
76. **Building Trust: Focusing on trust to improve health care**. ABIM Foundation. Accessed November 9, 2023. <https://abimfoundation.org/what-we-do/rebuilding-trust-in-health-care>.
77. **Carraccio C, Lentz A, Schumacher DJ**. “Dismantling Fixed Time, Variable Outcome Education: Abandoning ‘Ready or Not, Here they Come’ is Overdue.” *Perspect Med Educ*. 2023; 12(1): 68–75. DOI: <https://doi.org/10.5334/pme.10>

TO CITE THIS ARTICLE:

Caretta-Weyer HA, Smirnova A, Barone MA, Frank JR, Hernandez-Boussard T, Levinson D, Lombarts KMJMH, Lomis KD, Martini A, Schumacher DJ, Turner DA, Schuh A. The Next Era of Assessment: Building a Trustworthy Assessment System. *Perspectives on Medical Education*. 2024; 13(1): 12–23. DOI: <https://doi.org/10.5334/pme.1110>

Submitted: 16 July 2023 **Accepted:** 18 December 2023 **Published:** 22 January 2024

COPYRIGHT:

© 2024 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Perspectives on Medical Education is a peer-reviewed open access journal published by Ubiquity Press.