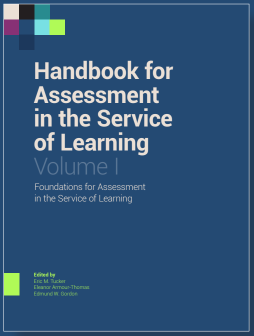
Looking Back, Moving Forward: Reflections on the Foundations for Assessment in the Service of Learning

Eleanor Armour-Thomas, Sheryl L. Gómez, and Eric M. Tucker

UMassAmherst
University Libraries



Series Editors:
Edmund W. Gordon, Stephen G. Sireci, Eleanor Armour-Thomas, Eva L. Baker, Howard T. Everson, and Eric M. Tucker

ISBN: 978-1-945764-33-2

Suggested Citation:
Armour-Thomas, E., Gómez, S. L., & Tucker, E. M. (2025). Looking back, moving forward: Reflections on the foundations for assessment in the service of learning. In E. M. Tucker, E. Armour-Thomas, & E. W. Gordon (Eds.), *Handbook for assessment in the service of learning, Volume I: Foundations for assessment in the service of learning.* University of Massachusetts Amherst Libraries.

Looking Back, Moving Forward: Reflections on the Foundations for Assessment in the Service of Learning

Eleanor Armour-Thomas, Sheryl L. Gómez, and Eric M. Tucker

Volume I invites us to approach assessment as integral to pedagogy—a continuous dialogue that guides learning—so it functions as an instrument for understanding—and not only as an instrument for determining status. Professor Edmund W. Gordon's charge, in this regard, is clear: rebalance assessment's role and embrace it as an approach for cultivating intellective competence—as a relational, pedagogical enterprise (Gordon, 2025). We must measure what matters—valuing transferable competencies; innovate how we assess—using authentic, interactive tasks grounded in how people learn; and measure in a manner that is useful and usable while upholding technical requirements for validity, fairness, scientific soundness, transparency and credibility.

Designing Assessment in the Service of Learning: Principles and Frameworks

Section 1 argues that rebalancing the focus of assessment begins with principled design. What we intend to infer must govern how evidence is generated, scored, and interpreted (AERA, APA, & NCME, 2014; Pellegrino, Chudowsky, & Glaser, 2001). Grounded in how people learn, principled design aligns claims, observations, and interpretations, embeds tasks in instruction, and meets heightened obligations for transparency, reliability, and fairness (Darling-Hammond & Adamson, 2014). Section 1 offers architecture as well as aspiration: a coherent infrastructure linking cognitive models, measurement, and use, including while designing and developing performance task-, simulation-, game-, or AI-powered assessment.

Grounding Assessment in How People Learn: Scientific Foundations

Section 2 grounds assessment firmly in contemporary science of learning and development: to serve learning, assessment must track understanding as it actually unfolds—dynamic, relational, and situated—and return useful and usable evidence to educators and learners (Nasir, Lee, Pea, & McKinney de Royston, 2020). It adopts an individual-in-context stance, recognizing that cognition grows through interactions with social, cultural, and material environments. The pedagogical throughline is clear: evidence should arise inside instruction, through discipline-specific tasks that make thinking visible and provide timely feedback so teachers can adjust and students can take the next step (Black & Wiliam, 1998). Efficacy is designed from the start by accommodating and recognizing the strengths of human variation and offering multiple, authentic ways to demonstrate competence (Armour-Thomas, McCallister, Boykin, & Gordon, 2019). The section advances practical measurement for improvement—frequent, purpose-built checks that monitor growth over time—and argues for coherence and balance across levels of the education system (LeMahieu& Cobb, 2025; Marion, Pellegrino & Berman, 2024).

Harnessing Emerging Innovations and New Possibilities

Section 3 adopts critical optimism about emerging technologies, arguing that innovation must be purpose-first and human-centered. Artificial intelligence, simulations, and game-based environments can have the affordances to elicit richer evidence of complex skills, personalize challenge and support, and deliver timely feedback, making assessment a seamless part of learning rather than a separate event. Such tools demand strong validity arguments, vigilance for bias, transparent operation, and firm commitments to privacy, data stewardship, and fair access.

The Call to Operationalize Pedagogical Analysis

Through a new paradigm Gordon calls Pedagogical Analysis, he champions the importance of collecting and analyzing pedagogical evidence—complementing status measures—so we can understand and improve assessment–teaching–learning transactions. Status evidence remains vital, but on its own is insufficient to “inform and improve teaching and learning processes and outcomes” (Gordon Commission, 2013, p. vii). that catalyze and cultivate intellective competence in all learning persons.

Conclusion: From Foundations to Frameworks to Transformation

Volume I seeks to understand assessment as principled design and a relational act whose purpose is to advance teaching and learning. This means building systems that do more than register outcomes—they must illuminate how students think, struggle, and progress, so that teaching with supports can be responsive to their assets, needs and interests. Volume II advances the 'how,' bridging design principles and technical innovation; Volume III demonstrates those principles in action across varied contexts. If we measure what matters, innovate how, and measure well, assessment becomes an ally of teaching and a lever for learning—pursuing Gordon's vision of assessment in the service of learning.

References

(AERA, APA, NCME) American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Armour-Thomas, E., McCallister, C., Boykin, A. W., & Gordon, E. W. (Eds.). (2019). *Human variance and assessment for learning*. Third World Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74.

Darling-Hammond, L., & Adamson, F. (2014). *Beyond the bubble test: How performance assessments support 21st-century learning*. Jossey-Bass.

Gordon, E. W. (2025). Series introduction: Toward assessment in the service of learning. In E. M. Tucker, E. Armour-Thomas, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning, Volume I: Foundations for Assessment in the Service of Learning.* University of Massachusetts Amherst Libraries.

LeMahieu, P., & Cobb, P. (2025). Practical Measurement for Improvement: Foundations, Design, Rigor. In E. M. Tucker, E. Armour-Thomas, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning, Volume I: Foundations for Assessment in the Service of Learning.* University of Massachusetts Amherst Libraries.

Marion, S. F., Pellegrino, J. W., & Berman, A. I. (Eds.). (2024). *Reimagining balanced assessment systems.* National Academy of Education.

Nasir, N. S., Lee, C. D., Pea, R., & McKinney de Royston, M. (Eds.). (2020). *Handbook of the cultural foundations of learning*. Routledge.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press. https://doi.org/10.17226/10019

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.