

From Aspiration to Application: Working Examples of Assessment in the Service of Learning

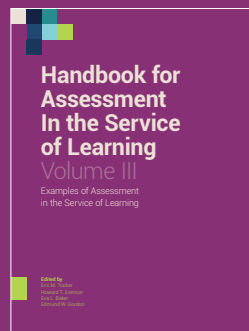
Eva L. Baker, Howard T. Everson,
and Eric M. Tucker

UMassAmherst

University Libraries

Series Editors:

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





© 2025 by Eva L. Baker, Howard T. Everson, and Eric M. Tucker

The Open Access version of this chapter is licensed under a Creative Commons Attribution–NonCommercial–NoDerivatives 4.0 International License (CC-BY-NC-ND 4.0).

ISBN: 978-1-945764-33-2

Suggested Citation:

Baker, E. L., Everson, H. T., & Tucker, E. M. (2025). From aspiration to application: Working examples of assessment in the service of learning. In E. M. Tucker, E. L. Baker, H. T. Everson, & E. W. Gordon (Eds.), *Handbook for assessment in the service of learning, Volume III: Examples of assessment in the service of learning*. University of Massachusetts Amherst Libraries.

From Aspiration to Application: Working Examples of Assessment in the Service of Learning

Eva L. Baker, Howard T. Everson, and Eric M. Tucker

This chapter has been made available under a CC BY-NC-ND license.

Building on the vision articulated in the Series Introduction (Gordon, 2025), this volume answers the call to bridge the chasm between the aspiration for assessment in the service of learning and its practical application. It moves from the 'why' to the tangible 'how' by presenting the 'actionable blueprints' Gómez (2014) called for: concrete examples of assessments that support learning. Drawn from contexts as varied as the College Board's AP® Art and Design portfolios and game-based assessments, these examples are aligned with the core design principles outlined in Baker, Everson, Tucker, and Gordon (2025).

A Framework for Analysis: Three Complementary Lenses

To provide context for these examples, we offer a framework of three complementary 'lenses' from the work of Robert J. Mislevy: Assessment as Evidentiary Argument, as a Feedback Loop, and as Social Practice (Mislevy, 2012, 2018; Bell & Mislevy, 2021). This three-part framework provides a lens for analyzing the working examples that follow, complementing the design principles for assessment in the service of learning proposed by Baker et al. (2025).

The sheer variety of the chapters that follow—from youth development programs to widely adopted digital courseware—calls for shared language for analysis. These examples do more than simply illustrate promising directions for assessment; they reveal aspects of the underlying architecture of learning-oriented assessment designs. To fully appreciate the design trade-offs and innovations detailed ahead, Mislevy's framework invites readers to move beyond viewing these chapters as simple narrations and instead engage with them as complex case studies in assessment design, analyzing how each exemplar succeeds, and where it faces challenges, in integrating the interdependent demands of valid evidence (argument), actionable feedback (loop), and authentic context (practice). This analytical approach is essential for synthesizing insights across chapters and understanding how each contributes to a broader vision of assessment in the service of learning.

Assessment as Evidentiary Argument

The first lens reframes assessment not as a simple measurement tool but as a structured, evidence-based argument (Pellegrino, Chudowsky, & Glaser, 2001; Kane, 2013; Mislevy, Steinberg, & Almond, 2003). From this perspective, a student's performance serves as the data used to support an inference or interpretive claim about their knowledge, skills, abilities, or other attributes. This connection is justified by a warrant and its backing (a generalization supported by theory), requiring designers to first articulate their claims and then construct tasks to elicit the necessary evidence to support those claims.

Assessment as a Feedback Loop

The second perspective shifts focus from the quality of evidence to its use, emphasizing that the data's value depends on how well it informs subsequent decisions. This logic, therefore, requires designers to consider who needs the assessment information, when, and in what form (Hattie & Timperley, 2007; Shute, 2008). It also exposes the tension between a teacher's need for immediate

instructional feedback (a focused, shorter loop) and a system leader's annual data needs (a wider, longer loop). Because an assessment optimized for one purpose is suboptimal for the other, this logic compels a move toward coherent systems of assessments, each designed for a specific purpose.

Assessment as Social Practice

The third lens allows for viewing assessment as a social and instructional activity. Drawing from a sociocognitive perspective, it recognizes that assessments are not neutral instruments but powerful cultural practices that signal what is valued and shape classroom interactions (Shepard, 2000; Bennett, 2023; Nasir, Lee, Pea, & McKinney de Royston, 2020; Penuel & Watkins, 2019). This logic pushes for authentic assessments that mirror real-world disciplinary practices, blurring the line between learning and assessing so that the assessment itself becomes a meaningful learning experience (Mislevy, 2012; Bell & Mislevy, 2021). This perspective also brings issues of human variation and equity to the forefront (Gordon, 1995). It aligns with Gordon's (2020) assertion that designing assessments to respect learners' varied backgrounds and cultivate their abilities is a moral and civil rights imperative. This imperative is a through-line in the chapters that follow, which feature assessments designed for a broad range of learners, from young children interacting with educational media to middle years students developing foundational reading skills.

The Integrated Architecture of Learning-Oriented Assessment

These three perspectives are complementary not separate; together they define the architecture of learning-oriented assessment. The exemplars in this volume show that the promise of innovation rests not primarily on emerging technology, but on the thoughtful integration of their forms of reasoning about assessments intended to support learning. An assessment's capacity to improve learning depends on its ability to elicit valid evidence, provide useful and actionable feedback, and situate itself meaningfully in the social context of teaching and learning (Darling-Hammond, Herman, Pellegrino, Abedi, Aber, Baker, Bennett, Gordon, Haertel, Hakuta, Ho, Linn, Pearson, Popham, Resnick, Schoenfeld, Shavelson, Shepard, Shulman, & Steele, 2013; Goldman & Lee, 2024).

Conclusion

This volume's tangible examples, from badges as assessments to standards-aligned tests and assessments, are offered not as fully formed solutions but as invitations to reflect, iterate, and build upon. They provide the field with a set of powerful existence proofs, hopefully inspiring and better equipping test developers, researchers, and educators to construct more coherent, learner-centered assessment systems that genuinely promote learning and achievement for all learners.

References

- Baker, E. L., Everson, H. T., Tucker, E. M., & Gordon, E. W. (2025). Principles 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.
- Baker, E. L., & Gordon, E. W. (2014). From the assessment of education to the assessment for education: Policy and futures. *Teachers College Record*, 116, 1–24.
- Bell, C., & Mislevy, R. (2021). Practice, feedback, argument, measurement: A frame for understanding diverse perspectives on teaching assessments. In M. Blikstad-Balas, K. Klette, & M. Tengberg (Eds.), *Ways of analyzing teaching quality: Potentials and pitfalls* (pp. 21–52). Scandinavian University Press. <https://doi.org/10.18261/9788215045054-2021-01>
- Bennett, R. E. (2023). Toward a theory of socioculturally responsive assessment. *Educational Assessment*, 28(2), 83–104. <https://doi.org/10.1080/10627197.2023.2202312>
- Darling-Hammond, L., Herman, J., Pellegrino, J. W., Abedi, J., Aber, J. L., Baker, E., Bennett, R., Gordon, E., Haertel, E., Hakuta, K., Ho, A., Linn, R. L., Pearson, P. D., Popham, J., Resnick, L., Schoenfeld, A. H., Shavelson, R., Shepard, L. A., Shulman, L., & Steele, C. M. (2013). *Criteria for high-quality assessment* [Technical Report]. Stanford Center for Opportunity Policy in Education.
- Goldman, S. R., & Lee, C. D. (2024). Human learning and development: Theoretical perspectives to inform assessment systems. In S. F. Marion, J. W. Pellegrino, & A. I. Berman (Eds.), *Reimagining balanced assessment systems* (pp. 48–92). National Academy of Education.
- Gómez, L. M. (2014). The Gordon Commission: An opportunity to reflect. *Teachers College Record*, 116, Article 110301.

- The Gordon Commission on the Future of Assessment in Education. (2013). *To assess, to teach, to learn: A vision for the future of assessment* [Technical Report]. ETS.
https://www.ets.org/Media/Research/pdf/gordon_commission_technical_report.pdf
- Gordon, E. W. (1995). Toward an equitable system of educational assessment. *The Journal of Negro Education*, 64(3), 360–372.
- Gordon, E. W. (2020). Toward assessment in the service of learning. *Educational Measurement: Issues and Practice*, 39(3), 72–78.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73.
- Mislevy, R. J. (2012). *Four metaphors we need to understand assessment*. (Commissioned paper for The Gordon Commission on the Future of Assessment in Education). Educational Testing Service.
- Mislevy, R. J. (2018). *Sociocognitive foundations of educational measurement*. Routledge.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspectives*, 1, 3–67.
- 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 Academies Press. <https://doi.org/10.17226/10019>
- Penuel, W. R., & Watkins, D. A. (2019). Assessment to promote equity and epistemic justice: A use-case of a research-practice partnership in science education. *The Annals of the American Academy of Political and Social Science*, 683(1), 201–221. <https://doi.org/10.1177/0002716219843249>

- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.