

Reconceptualizing Assessment to Improve Learning

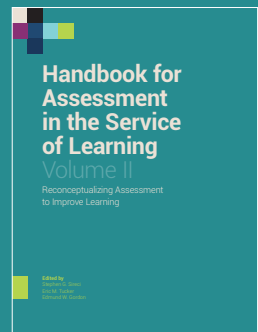
Eric M. Tucker and Stephen G. Sireci

UMassAmherst

University Libraries

Series Editors:

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





© 2025 by Eric M. Tucker and Stephen G. Sireci

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:

Tucker, E. M., & Sireci, S. G. (2025). Reconceptualizing assessment to improve learning. In S. G. Sireci, E. M. Tucker, & E. W. Gordon (Eds.), *Handbook for assessment in the service of learning, Volume II: Reconceptualizing assessment to improve learning*. University of Massachusetts Amherst Libraries.

Reconceptualizing Assessment to Improve Learning

Eric M. Tucker and Stephen G. Sireci

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

"Never forget the world of the possible." I wrote down this quote from my friend and mentor, Professor Edmund W. Gordon, on February 23, 2024. It was about three years after he first invited me to work with him on a "Handbook for Assessment in the Service of Learning." Of course, I agreed to co-edit the Handbook—little did I envision just how far that initial idea would expand. What began as a plan for a single handbook has blossomed into a full series of Handbooks on Assessment in the Service of Learning. I am proud to have joined an all-star editorial team in bringing forth Volume II of this series. This journey has shown me *the world of the possible* in educational assessment, a world I would not have imagined just four years ago.

Stephen G. Sireci

Northampton, MA

We are proud to offer Volume II of the *Handbook for Assessment in the Service of Learning*, to all those who strive to help others through education. This volume, entitled *Reconceptualizing Assessment to Improve Learning*, serves a special role in this series. Volume I of the Handbook explored the foundational design principles and research bases for transforming assessment to inform teaching and learning processes, essentially making the case for why change is needed and outlining key design imperatives. Volume III, at the other end of the arc, will illustrate practical implementations and working examples—case exemplars of assessment approaches that embody aspects of these new approaches. Volume II stands as the conceptual and methodological bridge between these two. In these pages, we move from Volume I's focus on research, design, and technology to reconceptualization and innovation, redefining what assessment can be and providing prototypes of how to do it. Our focus here is on abandoning outdated traditions of educational testing in favor of approaches to assessment that serve teachers and learners first. The chapters assembled in Volume II—contributed by an all-star cast of authors at the cutting edge of the field—offer examples of how we can design assessments that *truly* support learning: how we can harness new technologies to improve assessment, ensure our assessments meet the needs of all learners, and provide richer information for students, educators, and other stakeholders invested in education. In short, this volume tackles *how* we might reconceptualize assessment to fulfill the ambitious vision articulated by Professor Edmund W. Gordon and the Gordon Commission over a decade ago. The Gordon Commission on the Future of Assessment in Education (2013) argued that traditional testing—fixated on ranking students and certifying “what is”—must give way to approaches that illuminate how learning happens and how it can be improved. Our task in Volume II is to build on that vision, providing aspects of both the conceptual blueprints and the inventive tools needed to reinvent assessment in the service of learning.

To organize this rich body of work, Volume II is divided into two sections. Section I: Foundations and Frameworks for Learner-Centered Assessment lays out the key theories, principles, and frameworks guiding the transformation of assessment. These chapters articulate why we must reconceptualize assessment and on what bases—from formative assessment foundations to considerations of validity and social justice. Section II: Innovations in Practice—Tools and Methods Serving Learning then showcases a variety of cutting-edge approaches that put those

principles into action. The chapters in Section II present novel methodologies and tools—from game-based assessments and learner-centered portfolios to culturally responsive co-design and new uses of data—each illustrating how assessment can be embedded into educational practice to engage learners and provide meaningful feedback for improvement. In essence, Section I gives us the “why” and “what” of reconceptualizing assessment, while Section II explores some of the practical “how”—mirroring our series’ progression from Volume I’s foundational research and design approaches to Volume III’s applied cases.

Section I: Foundations and Frameworks for Learner-Centered Assessment

Section I curates aspects of the conceptual underpinnings for “assessment in the service of learning.” The six chapters in this section establish core approaches and big-picture ideas that set the stage for reimagining assessment as a tool to inform and improve learning, rather than merely to audit what Professor Gordon might call achieved intellectual competence. These chapters span formative assessment, self-regulated learning, personalization and equity, and theoretical frameworks for validity and justice—and together provide a meaningful contribution to a foundation for a learner-centered assessment approach.

- Susan M. Brookhart: *Developing Educational Assessments to Serve Learners*:** Susan Brookhart (2025) provides a perfect beginning for this volume. She reminds us that learning begins long *before* children enter formal schooling, and that formative assessment is the solid foundation on which all assessments in the service of learning are built. Brookhart identifies key factors needed to facilitate assessment for learning—a supportive learning culture, clear learning goals, and clear success criteria—underscoring what must be in place for an assessment to successfully serve learners. She follows these insights with practical guidance on creating assessments that yield the feedback students need to advance their learning. Her chapter reinforces the notion that a formative, feedback-rich culture is foundational to any effective learner-centered assessment system.
- Héfer Bembenutty: *Toward a Culturally Self-Regulated Dynamic Pedagogy Assessment System*:** In the next chapter, Héfer Bembenutty (2025) continues the theme of integrating curriculum, instruction, and assessment to support self-regulated learning. Extending Armour-Thomas and Gordon’s dynamic pedagogy framework, Bembenutty describes an approach that values students’

cultures and empowers learners to use assessment information to understand and guide their own learning (a model he refers to as *Culturally Self-Regulated Dynamic Pedagogy*). The assessment-pedagogy practices outlined in this chapter go beyond simply adding formative assessments into instruction—they emphasize creating inclusive learning environments where assessment is woven into the learning process and students are active agents in their learning. By demonstrating how teaching, learning, and assessment processes can jointly foster students' self-regulation skills, this chapter exemplifies the deep integration of assessment with instruction to benefit learners.

- **Randy E. Bennett, Eva L. Baker, and Edmund W. Gordon: *Personalizing Assessment for the Advancement of Equity and Learning*:** Bennett, Baker, and Gordon (2025) also highlight culturally responsive principles and the importance of learner variation. This chapter illustrates how assessment in the service of learning can be designed to advance equity by personalizing the assessment process. The authors propose using personalized assessments to accommodate the wide range of variation in the learner population. They review research on learner variability and describe different conceptualizations of diversity, then offer concrete principles for adapting both learning activities and assessments to build on learners' individual experiences, cultures, and identities. These principles form a helpful roadmap for designing assessments that can flexibly meet the needs of diverse learners, ensuring that assessment practices contribute to equity and do not treat fairness and responsiveness as an afterthought.
- **Norris M. Haynes, Mary K. Boudreaux, and Edmund W. Gordon: *A Theory-Informed and Student-Centered Framework for Comprehensive Educational Assessment*:** Haynes, Boudreaux, and Gordon (2025) present a broad theoretical framework to guide learner-centered assessment. They draw on three major perspectives—constructivism, sociocultural theory, and implicit theory—to ensure that assessments for learning provide valid insights into *how* students learn and develop. By acknowledging the influence of school and classroom culture and climate (including the constraints of the “hidden curriculum”), this chapter shows how assessments can make more meaningful learning experiences that support the holistic development of all learners. The authors discuss a comprehensive range of assessment types (formative, summative, diagnostic, ipsative, self-assessment, norm- and criterion-referenced, curriculum-based, etc.), illustrating how each can be employed,

referenced, curriculum-based, etc.), illustrating how each can be employed, in line with sound learning theory, to support student growth. This wealth of approaches, grounded in the learning sciences, enriches our toolbox for designing assessments that are both rigorous and learner-centered.

- **Stephen G. Sireci and Danielle M. Crabtree: *Validity Theory and Validation of Assessments in the Service of Learning*:** In this chapter, Sireci and Crabtree (2025) tackle questions of validity for assessments whose primary purpose is to serve learning. They explain traditional notions of test validity and how classic validity theory applies when assessments are repurposed to support learning. They discuss how to gather and evaluate validity evidence to ensure assessments in the service of learning are actually accomplishing their intended goals. The chapter aims to rectify the glaring lack of practitioner (teacher) perspectives missing from many test development and validation processes. Their chapter draws largely on the established standards of educational testing (e.g., American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), but also introduces newer perspectives tailored to the specific goals of using assessment to enhance learning for students. Ultimately, they argue that transforming assessment does not mean abandoning rigor; rather, it requires expanding our concept of rigor to include usefulness for learning. Assessments in the service of learning must meet high standards of technical quality and yield information that is instructionally actionable and meaningful to students and teachers.
- **Stephen G. Sireci, Sergio D. Araneda, and Kimberly A. McIntee: *Social Justice in Educational Assessment: A Blueprint for the Future*:** Sireci, Araneda, and McIntee (2025) round out Section I with a chapter that poses a provocative, forward-looking question: *How can assessment be reimaged as a tool for social justice in education?* In this concluding chapter of Section I, they argue that issues of fairness, equity, and justice should not be afterthoughts in assessment design—they must be treated as foundational principles from the very start. They examine ways in which current assessment practices can unintentionally perpetuate inequities (for instance, through cultural bias in test content, unequal access to test preparation, or high-stakes uses that disproportionately impact marginalized groups). They also outline strategies to ensure assessment systems promote equity and empowerment rather than

reinforce disparities. These strategies include designing culturally responsive assessments that value diverse ways of knowing, involving students and communities in co-designing assessments, and using assessment data proactively to identify and close opportunity gaps (instead of using data punitively to label or punish). In essence, this chapter elevates the conversation to the policy and ethical level, extending the learner-centered assessment narrative into the realm of social responsibility. They contend that a truly learning-centric assessment system must also be a justice-centric system—one that actively works to dismantle historical biases and create more inclusive, supportive educational environments. By articulating concrete principles and recommendations for socially just assessment, they provide both a moral compass and a practical guide for the future. This chapter challenges us to ensure the transformation of assessment remains aligned with the broader goal of educational equity.

Although each of these chapters has a distinct focus, together they offer components of an emerging framework for rethinking assessment. They collectively prompt us to reconsider *why* we assess (our purposes), *what* we assess (the constructs and competencies we value), and *how* assessment practices impact learners and society (the consequences we care about). In doing so, Section I lays robust groundwork for reinventing assessment as a positive force in the learning process. Volume I gave us the vision and the design imperatives for assessment reform, and here in Volume II—especially through Section I—we undertake the critical work of reconceptualization, redefining the fundamental ideas and frameworks on which future assessments will be built. These conceptual foundations now pave the way for the innovations presented in Section II, where theory meets practice.

Section II: Innovations in Practice—Tools and Methods Serving Learning

If Section I explains why and on what insights we must change assessment, Section II explores important aspects of how those principles can be realized through new approaches and tools. The chapters in Section II showcase a range of innovative practices that embed assessment into the fabric of teaching and learning. From games and portfolios to data analytics and co-designed assessments, each contribution breathes life into the learner-centered vision by demonstrating concrete strategies for making assessment an integral, engaging

part of education. Collectively, these chapters show that the lofty ideals outlined in Section I can indeed be translated into inventive designs that improve learning.

- **James Paul Gee: *Game-Based Learning: A Design-Based Theory of Teaching–Learning–Assessment Systems*:** Jim Gee (2025) opens Section II with a fascinating illustration of how assessment can be seamlessly integrated with instruction to foster engagement and deep learning. He presents a design-based theory of “teaching–learning–assessment” systems grounded in what we can learn from good games. As Gee insightfully observes, “good games are good for teaching, learning, and assessment”. In a well-designed game, a player’s constant problem-solving and immediate feedback naturally generate evidence of learning; the assessment is essentially woven into the gameplay itself. Gee’s chapter explains how games can integrate teaching, learning, and assessment invisibly (to the learner) yet effectively, and he provides a blueprint for developing such game-based assessment systems. Beyond theory, this chapter offers practical design principles for educators and developers interested in creating engaging, game-like assessments that motivate learners and simultaneously yield rich information about their learning processes.
- **Carol A. Bowman and Edmund W. Gordon: *The Educative/Learning Portfolio: Towards Educative Assessment in the Service of Human Learning*:** Bowman and Gordon (2025) reintroduce a more familiar, yet underutilized, assessment tool—the portfolio—and reconceptualize it as an “educative portfolio.” They describe how a student’s portfolio of work can be transformed from a static showcase of accomplishments into a dynamic process and instructional tool that actively cultivates learning. In an *educative portfolio* model, compiling and reflecting on one’s work becomes an integral part of the learning experience itself. Students select pieces, reflect on their growth, and discuss their work, meaning the assessment happens through those activities. This approach has the potential to yield rich, authentic evidence of learning—in fact, the portfolio artifacts provide “more useful and abundant evidence of achievement than a simple metric,” offering a revealing window into the processes of the student’s learning. Unlike the “stealth” assessment in a game, portfolio-based assessment is purposefully visible and transparent: clear objectives, expectations, and reflective actions before, during, and after the assessment are central. Bowman and Gordon show how transparency and reflection in portfolio assessment support learning, making the process educative for

the student. Like Gee's games, the portfolio chapter exemplifies integrating assessment with curriculum and instruction—albeit in a different form—and demonstrates that even traditional assessment formats can be innovated to serve learning more effectively.

- **Maria Elena Oliveri, Kerrie A. Douglas, and Mya Poe: *Building Culturally and Linguistically Responsive Workplace Assessments for Learning*:** Oliveri, Douglas, and Poe (2025) advance the idea of culturally responsive assessment through the lens of workplace learning. They illustrate how involving learners (and other stakeholders) directly in the test development process—for instance, via participatory co-design—leads to assessments that are more valid and appropriate for diverse populations. After introducing the concept of culturally and linguistically responsive teaching and assessment, the authors present a compelling use case from engineering education. In this example, assessments were co-designed to reflect multilingual, multicultural workplace realities. The chapter demonstrates that when assessments are grounded in learners' cultural contexts and allow multiple ways for learners to demonstrate competence, the assessments become not only fairer but also more instructionally valuable. In the context of workplace learning, this means assessments better prepare and reflect what learners need on the job, while honoring the diverse backgrounds they bring. Oliveri, Douglas, and Poe offer practical guidance for developing assessments *with* learners rather than *for* learners, embodying the principle that assessment design should adapt to learners (instead of expecting learners to adapt to rigid assessments). This culturally responsive co-design approach demonstrates how we can develop assessments that genuinely include and empower every learner.
- **Stephen G. Sireci and Neal Kingston: *Removing the "Psycho" from Education Metrics*:** In this provocatively titled chapter, Sireci and Kingston (2025) examine aspects of how assessment results are reported and used. They critique traditional testing metrics and reporting formats, which too often mystify or alienate educators and students by drowning them in psychometric complexity. Their chapter advocates for reimagining how assessment results are communicated to serve learning, and they argue for shifting from obscure statistics to intuitive, learner-centered feedback that students, teachers, and parents can readily understand and act upon. As a working example, they describe the *Dynamic Learning Maps* system—an innovative assessment

designed for students with significant cognitive disabilities—as an illustration of assessment geared toward diagnosing individual learning needs and guiding instruction, rather than merely cataloging deficits. This system aims to provide rich diagnostic profiles of what a student *can do* and what might help them progress next, exemplifying assessment as a supportive tool for learning. Throughout the chapter, they show how tests and score reports can be designed in plain, user-friendly language without sacrificing the depth of information. By redesigning score reports to emphasize clear, actionable insights (what skills a student has mastered, and what they should work on next), they illustrate how to maintain rigor while making data more useful. The message across this chapter (and Oliveri et al.'s as well) is that assessments can and should adapt to learners and educators—not the other way around—by providing information that is accessible, meaningful, and geared toward helping every student learn.

- Gregory K. W. K. Chung, Tianying Feng, and Elizabeth J. K. H. Redman:**
Using Learner-System Interactions as Evidence of Student Learning and Performance: Chung, Feng, and Redman (2025) push the frontier of assessment by asking: What if every interaction a learner has with educational materials could count as assessment data? In this final chapter of Section II, they explore how emerging technologies and data analytics enable entirely new forms of evidence of student learning. The authors posit that every click, response, or choice a student makes in a digital learning environment is potential data about their thinking and skills. By capturing these fine-grained learner–system interactions, for example, how a child approaches problems in an online math game, we can glean insights that no traditional test alone could offer. Chung et al. outline methods for identifying which learner behaviors to capture, how to record them, and how to analyze this flood of data to draw valid inferences about student learning. Their approach applies rigorous measurement principles to forms of behavioral data not typically considered in assessment. They show how students' interactions with digital tutoring systems, educational games, and other instructional software can be interpreted as assessment evidence and modeled to provide ongoing feedback. The chapter includes vivid examples, such as observing preschoolers' strategies in a math game, to illustrate how these interaction data can reveal learning processes and guide instruction in real time. Although the context of their example is early childhood mathematics, the underlying

principles generalize to all levels and subjects: modern technology allows us to embed assessment into virtually any learning activity. This work highlights how assessment is evolving into something much broader than tests—it can encompass the *continuous stream of data* generated by learners as they engage with learning materials.

Each of these chapters advances understanding of practical approaches to make assessment more integrated with learning. Taken together, the contributions in Section II span a remarkable range of contexts and methods—but they all show how the core ideals from Section I can be realized in practice. Whether through immersive games, reflective portfolios, co-designed culturally responsive tasks, reimagined score reports, or data-rich digital environments, these authors are breathing life into the idea of assessment as a tool for learning. They exemplify the creativity and dedication needed to turn assessment from a once-a-year audit into an ongoing, student-centered conversation about growth.

Emergent Themes Across Volume II

Across both sections of Volume II, several key themes reverberate, weaving a unifying narrative of what it means to make assessment truly learner-centered:

- **Formative feedback and improvement:** A shift toward feedback-rich, formative practices (Shute, 2007) is evident throughout the volume. From Brookhart's emphasis on a formative culture to Bembenutty's focus on real-time self-regulation, the idea of using assessment to provide continuous feedback for improvement is a common thread. Even in Section II's tools, we see this theme: Gee's game-based assessments offer frequent feedback in context, Bowman and Gordon's portfolios embed feedback through reflection, and Chung et al.'s analytics turn interactions into actionable feedback. The notion that assessment should *inform* and *guide* learning—rather than merely judge it—underpins these chapters. (Gordon Commission on the Future of Assessment in Education, 2013; Gordon, 2020)
- **Learner agency and engagement:** The authors in this volume consistently treat learners as active participants in the assessment process, not passive subjects of measurement. This commitment to *learner agency* shows up in many forms: Bembenutty's culturally self-regulated pedagogy empowers students to monitor their learning, Oliveri et al.'s participatory co-design actively involves learners in creating assessments, and portfolio assessment

(Bowman & Gordon, 2025) gives students voice and choice in showcasing their learning. Even Gee's games put the learner in charge of problem-solving within the assessment environment. In line with Professor Gordon's vision (Cauce & Gordon, 2013), the student is not a mere object of assessment but an agent whose engagement, self-management, and self-reflection are integral to the process. By fostering agency, these chapters suggest that assessment can actually *motivate* and empower learners.

- **Validity, fairness, and social justice:** A strong imperative around quality measures and the advancement of justice runs through Volume II. Nearly every chapter grapples with how to make assessment more fair, inclusive, and beneficial for *all* learners. Bennett et al. explicitly center diversity by personalizing assessment to learner needs; Haynes et al. emphasize culturally attuned frameworks and holistic validity; Oliveri et al. design for linguistic and cultural responsiveness in diverse contexts; and our social justice chapter insists on making fairness and justice fundamental criteria for any assessment system. At the same time, maintaining validity and rigor is a shared concern—the volume does not advocate diminishing the importance of technical quality, but rather expanding our definitions of quality. Sireci and Crabtree's chapter, for example, shows how we can uphold rigorous validation standards for new kinds of assessments, ensuring that innovative assessments yield trustworthy evidence about student learning *while* avoiding cultural bias or misuse. In sum, Volume II envisions assessment systems that are high quality, rigorous, and just—assessments that earn stakeholders' confidence through validity *and* demonstrate a commitment to fairness and social responsibility.
- **Dynamic integration of assessment, learning, and instruction:** A recurring theme is the blurring of lines between assessment and instruction. Many authors echo Professor Gordon's call to integrate assessment in the processes of instruction rather than treat it as a separate, after-the-fact event. (Armour-Thomas & Gordon, 2013, 2025) Brookhart and Bembenuddy set the stage by describing classroom cultures where assessment is part of everyday teaching and learning. In Section II, this integration becomes concrete: in Gee's chapter, assessment *is* the gameplay; in Bowman's, assessment is woven into the act of curating and reflecting on learner work; in Chung et al's, assessment data is captured as students learn in digital environments. The benefit of such integration is twofold: it makes assessment more natural and less

anxiety-provoking, and it aims to produce more instructionally relevant data. Throughout the volume, we see that integrating assessment with instruction has the potential to lead to more timely insights and create a more supportive experience for learners—fulfilling the ideal of assessment as a “pedagogical transaction” embedded in learning.

- Responsiveness and relevance: Finally, responsiveness and the cultural foundations of learning emerge as a vital theme (Bennett et al., 2024; Mislevy et al., 2024; Nasir et al., 2020). The chapters collectively recognize that learners bring significant variation in backgrounds, languages, and ways of knowing to the table, and that assessments must honor and reflect that variation. This is most explicit in works like Oliveri et al.'s co-designed assessments for multicultural settings and Bennett et al.'s personalized approaches for diverse students. Haynes et al. also incorporate sociocultural perspectives to ensure assessments are meaningful across different contexts, and our social justice chapter takes this further to address systemic biases. Even outside the equity-focused chapters, cultural relevance appears in Gee's attention to engaging all learners through game narratives and in portfolio assessment's accommodation of individual expression. The through-line is that one-size-fits-all assessments are no longer acceptable; to truly serve learning, assessment practices must be adaptable to cultural and individual variation. By making assessments more responsive to learners' contexts, we not only improve fairness but also make assessment results more meaningful and actionable for each learner.

These themes—formative feedback, learner agency, validity and fairness, dynamic integration with instruction, and cultural responsiveness—resonate throughout portions of Volume II and tie the chapters together. They reflect a shared commitment to redefining assessment as something fundamentally in service of learning and human development.

Looking Ahead to Volume III

As we begin Volume II, it is worth reflecting on how the insights gathered here might set the stage for the third volume in our series. Volume III will carry this work forward by showcasing implementations and exemplars of assessment in the service of learning. In Volume III, readers will see the concepts and innovations from Volumes I and II come alive in various teaching and learning contexts. We

will explore case studies and models of assessment systems that have been implemented, demonstrating the impact and feasibility of the ideas we've been discussing. In a sense, if Volume II provides frameworks and tools, Volume III will show some functioning working examples—works in progress—built upon those designs principles.

The chapters of Volume II provide key tools in a conceptual and methodological toolkit for readers seeking to transform assessment. They have illustrated aspects of “the world of the possible” that was beyond our own vision when this project began. Now, Volume III will challenge us to apply that toolkit and learn from concrete experiences. By connecting theory to practice, Volume III will complete the bridge that Volume II has built between foundational principles and practical realization. We are excited to see how innovative assessments *in action* can further validate these ideas, reveal new challenges, and inspire continued refinement of assessment for learning.

In closing, we feel a deep sense of gratitude. First, we want to thank the authors of these chapters, who are pioneers in our field—their dedication and creativity made this volume possible. We are grateful as well to our fellow editors and collaborators; working with an editorial team of such vision and expertise has been a privilege. Above all, we thank Professor Edmund W. Gordon—the Professor—whose unwavering vision has guided this journey from the start. Over twelve years ago, Professor Gordon wrote of the coming “conflict and contradiction” between traditional assessment practices and new scientific developments. He challenged us to resolve that conflict by reimagining assessment in bold ways. This Handbook series is, in many ways, one response to that challenge. As Gordon and the Gordon Commission foresaw, science, technology, and imagination have opened new possibilities for assessment. The work in Volume II represents one collective effort to advance understanding of how we might realize those possibilities—to redesign assessment in light of what is now *possible* for the betterment of learners.

With profound gratitude and optimism, we invite you to engage with the chapters of this volume. We hope you find in them not only rigorous scholarship and practical insight, but also the same sense of hope and inspiration that we have found. Together, may we continue to explore and expand *the world of the possible* in assessment, in service of every learner's growth.

References

- 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., & Gordon, E. W. (2013). *Toward an understanding of assessment as a dynamic component of pedagogy*. Educational Testing Service.
https://www.ets.org/Media/Research/pdf/armour_thomas_gordon_understanding_assessment.pdf
- Armour-Thomas, E., & Gordon, E. W. (2025). *Principles of dynamic pedagogy: An integrative model of curriculum, instruction, and assessment for prospective and in-service teachers*. Routledge.
- Bembenutty, H. (2025). Toward a culturally self-regulated dynamic pedagogy assessment system. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Bennett, R. E., Baker, E. L., & Gordon, E. W. (2025). Personalizing assessment for the advancement of equity and learning. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Bennett, R. E., Darling-Hammond, L., & Badrinarayan, A. (Eds.). (2024). *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy*. Routledge.
- Bowman, C. A., & Gordon, E. W. (2025). The educative/learning portfolio: Towards educative assessment in the service of human learning. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Brookhart, S. M. (2025). Developing educational assessments to serve learners. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.

- Cauce, A. M., & Gordon, E. W. (2013). *Toward the measurement of human agency and the disposition to express it*. Educational Testing Service. https://www.ets.org/Media/Research/pdf/cauce_gordon_measurement_human_agency.pdf
- Chung, G. K. W. K., Feng, T., & Redman, E. J. K. H. (2025). Using learner-system interactions as evidence of student learning and performance. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Gee, J. P. (2025). Game-based learning: A design-based theory of teaching–learning–assessment systems. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- 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*. Educational Testing Service.
- Gordon, E. W. (2020). Toward assessment in the service of learning. *Educational Measurement: Issues and Practice*, 39(3), 72–78.
- Haynes, N. M., Boudreaux, M. F., & Gordon, E. W. (2025). A theory-informed and student-centered framework for comprehensive educational assessment. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Mislevy, R. J., Oliveri, M. E., Slomp, D., Crop Eared Wolf, A., & Elliot, N. (2024). An evidentiary-reasoning lens for socioculturally responsive assessment. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 151–167). Routledge.
- Nasir, N. S., Lee, C. D., Pea, R. D., & McKinney de Royston, M. (Eds.). (2020). *Handbook of the cultural foundations of learning*. Routledge.

- Oliveri, M. E., Douglas, K., & Poe, M. (2025). Building culturally and linguistically responsive workplace assessments for learning. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Sireci, S. G., Araneda, S. D., & McIntee, K. A. (2025). Social justice in educational assessment: A blueprint for the future. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Sireci, S. G., & Crabtree, D. M. (2025). Validity theory and validation of assessments in the service of learning. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Sireci, S. G., & Kingston, N. (2025). Removing the "psycho" from education metrics. In S. Sireci, E. Tucker, & E. W. Gordon (Eds.), *Handbook for Assessment in the Service of Learning: Volume II*. University of Massachusetts Amherst Libraries.
- Shute, V. J. (2007). *Focus on formative feedback*. Educational Testing Service.
https://www.ets.org/research/policy_research_reports/publications/report/2007/hslv.html