

Perspectives on Socioculturally Responsive Assessment in Large-Scale Systems:

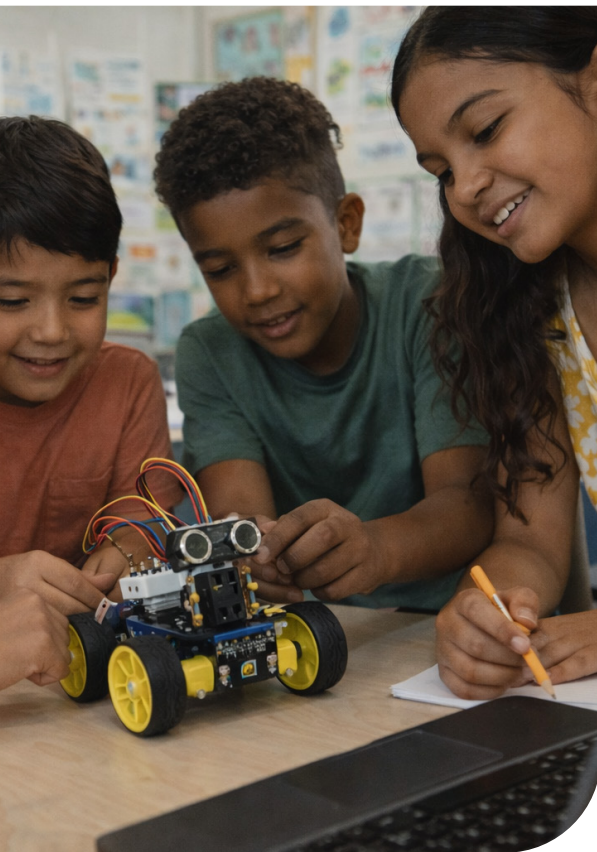
A Chapter Summary

Aneesha Badrinarayan, Randy E. Bennett, and
Linda Darling-Hammond



Perspectives on Socioculturally Responsive Assessment in Large-Scale Systems: A Chapter Summary

Aneesha Badrinarayan, Randy E. Bennett, and Linda Darling-Hammond



Abstract

This summary synthesizes key insights from *Socioculturally Responsive Assessment: Implications for Theory, Measurement, and Systems-Level Policy* (Bennett et al., 2025), offering a cross-cutting analysis of how large-scale assessments can become more valid, equitable, and educationally meaningful when designed through a socioculturally responsive lens. The authors highlight four central themes: aligning assessment content with students' lived experiences; increasing personalization of assessments via testing processes; broadening construct definitions to reflect diverse cultural, linguistic, and epistemological perspectives; and using frameworks and tools that guide inclusive development and interpretation. Drawing on examples such as the KĀ'EO Hawaiian language immersion assessment, Smarter Balanced, and adult education programs, the chapter illustrates how sociocultural responsiveness can be integrated into both test content and administration while maintaining technical quality. It also surfaces enduring challenges, including tensions between comparability and local validity, the need for assessments to reflect community values, and the evolving role of technology in supporting adaptive and culturally grounded assessments. The authors conclude by naming critical open questions—such as whose culture is centered, how assessment use-cases influence design, and how to ensure assessments both reflect and serve diverse learners—that must be addressed to advance Socioculturally Responsive Assessment (SCRA) as a transformative and sustainable paradigm for large-scale educational assessment.

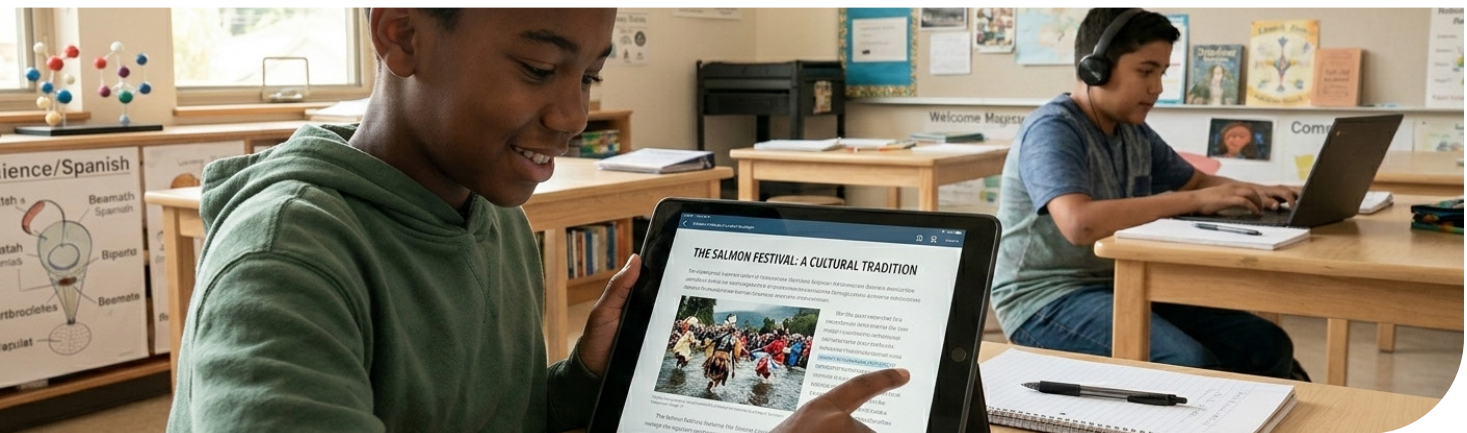
This summary has been made available under a CC BY-NC-ND license.

Introduction

At the heart of socioculturally responsive assessments (SCRA) are goals to intentionally account for the fact that learning—and demonstration of that learning—is inherently a social and cultural activity (Nasir et al., 2020; National Academies of Science, Engineering, and Medicine, 2018), and that assessments that do not account for how students develop knowledge cannot hope to accurately surface and communicate what a diverse range of students knows and can do. This fact influences both what we assess as well as how we surface evidence of learning.

In compiling the recent edited volume *Socioculturally Responsive Assessment: Implications for Theory, Measurement, and Systems-Level Policy* (Bennett et al. 2025d), we sought to synthesize a range of leading perspectives on approaches, technical considerations, and instrument-and-process designs intended to serve a common purpose: the development of assessments that are responsive to the unique and shared social, cultural, and linguistic experiences that shape how people learn and how they make what they know and can do visible. This summary offers insights drawn from across the range of scholarship presented in the volume to better characterize SCRA at scale, focused on emergent themes related to:

- 1 Relevance of assessment content
- 2 Personalization via assessment processes
- 3 Broadening of construct definition
- 4 Assessment development frameworks for SCRA



01

Theme One: The Importance of Relevance of Assessment Content to Students' Lived Experiences

Across many different conceptions, socioculturally responsive assessment design is predicated on the idea that how examinees interact with an assessment is not a fixed feature of assessment design, and instead is inextricably linked to the social, cultural, and linguistic contexts within which learning and performance happen. A goal, then, of SCRA is to bring assessment instruments into better alignment with examinees' experiences through various approaches to increasing relevance and personalization (Bennett 2023, 2024, 2025). Authors in the volume explore the relationship between students' funds of knowledge, test relevance, and performance through many different lenses, including disciplinary perspectives, approaches that attend to cultural ways of knowing, and ideas that account for relevance by broadening our conceptions of measurement targets.

Relevance Within Disciplinary Contexts: Insights From Reading Assessments

Reading assessments have been an area of focus for SCRA because of the considerable evidence that students' experience with and background knowledge of a given topic influences how they understand and make meaning of related texts. Text selection for large-scale assessments is a promising and important direction for content-based approaches to personalizing assessment instruments. Most large-scale reading assessments focus on common, provided texts as the basis for measuring students' abilities, and many qualitative and quantitative features of texts (e.g., text complexity, text types) can be accounted for while still allowing for more socioculturally grounded decisions about the nature of the texts students interact with on assessments.



Wang and colleagues (2025) offer empirical evidence to support the idea that using culturally relevant texts affects students' reading comprehension performance on large-scale assessments.

Skerrett and colleagues (2025) discuss how large-scale reading assessments can better account for the social nature of reading, students' funds of knowledge, and the range of student experiences with texts and topics. Drawing on the development of the 2026 National Assessment of Educational Progress Reading Framework, Skerrett et al. suggest several strategies for socioculturally responsive reading assessments, including:

- Using multimodal knowledge scaffolds that leverage video, audio, and graphical aids to support background knowledge students may need to access and make sense of the texts present on the exam;
- Allowing students to respond in home languages; and
- Embedding contextual probes (similar in many ways to the questions Ebe (2025) proposes in the Cultural Relevance Rubric) that can help assess student interest, motivation, and familiarity with texts.

Relevance to Specific Cultural and Linguistic Ways of Knowing.

A vital component of many conceptualizations of SCRA is the concept of cultural validity—the extent to which assessments account for the cultural, linguistic, and social repertoires students draw upon when interpreting and responding to disciplinary tasks (Solano-Flores & Ruiz-Primo, 2025; Nelson-Barber & Trumbull, 2025; Solano-Flores & Nelson-Barber, 2001). Cultural validity concerns how these repertoires shape meaning-making and influence not only performance but also the validity of inferences drawn about student understanding.

Nelson-Barber and Trumbull (2025), drawing on work with Indigenous communities, argue that cultural validity is as essential as construct validity: disregarding



students' cultural and linguistic contexts during assessment development and score interpretation introduces threats to overall validity. For example, for Indigenous students—including American Indian, Alaska Native, Native Hawaiian, and other Pacific Islander groups—histories, governance structures, belief systems, heritage languages, cultural values, and community-based practices are distinct from Western, Eurocentric norms. When assessments ignore these lived realities, they may:

- Present information in unfamiliar or incongruent ways that create barriers to comprehension;
- Require response modes that conflict with students' typical ways of demonstrating understanding;
- Apply narrow success criteria that devalue culturally rooted ways of knowing and sensemaking; and
- Reinforce experiences of dehumanization, marginalization, or erasure, which can undermine motivation and perseverance.

In all such cases, the result is the same: student performance no longer reflects what students know and can do, but rather their ability to navigate unfamiliar or invalidating assessment structures.

While these issues have been examined carefully in the assessment of Indigenous students, they also extend to many other populations whose cultural and linguistic experiences are underrepresented in mainstream testing assumptions.

Creating relevant large-scale assessments fundamentally requires that test developers generate content by making some assumptions about students—assumptions can range from extrapolating insights from deep and intentional student engagement to a broader set of students, to making assumptions that certain aspects of students' backgrounds (e.g., race, ethnicity) will confer experiences or perspectives to all students who share those elements of identity. The key ideas of choice on the part of the student and performance assessments that allow students to demonstrate their knowledge in broader ways are emphasized in our chapter as methods for allowing students to demonstrate their knowledge in culturally valid ways. These ideas are relevant to this and subsequent sections of this summary.

02

Theme Two: Personalizing Assessment Processes

Another approach to bringing assessment experiences into better alignment with students' social, cultural, and linguistic experiences is allowing for personalization or user-specific adaptations in ways students interact with assessments. Sireci and colleagues (2025) describe a framework for UNDERSTANDARDization in which decisions about what aspects of test administration are standardized are driven by learner-centered approaches that seek to better understand how different groups of students might understand and interact with the test, rather than rigid models of standardization that emphasize uniform testing conditions as a primary way to ensure fairness and validity of score interpretation. Sireci et al. (2025) discuss UNDERSTANDARDization in the context of two adult-centered assessments—the Adult Skills Assessment Program (ASAP) and the English Test for Adults (ETA)—both of which are intended to be administered to adults at scale.

Sireci and colleagues' UNDERSTANDARDization framework provides a useful model for examining administration-related approaches to SCRA. In the volume, authors describe two major approaches to this kind of personalization in K–12 assessment systems: (1) adaptations and accommodations that allow learners to better access the *same* content, and (2) allowing for student agency in deciding how students will demonstrate evidence of progress, proficiency, and mastery relative to common constructs via different content. For example, Michel and Shyyan (2025) describe the accessibility supports within the Smarter Balanced assessment system (SBAC) as a set of strategies for personalizing the testing experience such that all learners—particularly students with disabilities

and multilingual learners—can engage meaningfully with the same test content. Smarter Balanced assessments are built around relatively narrow constructs tied to states' math and English–Language Arts standards and are designed to be administered across diverse state contexts and student backgrounds. The assessments' primary goal is to generate evidence of student progress toward academic standards for federal accountability purposes within each state, which requires a certain amount of rigidity in content and administration.

Smarter Balanced provides multiple levels of support for personalization, including:

- Universal tools that are available to all students (e.g., digital notepad), allowing them to customize their testing experience to better account for how they make meaning and access content;
- Designated supports that are available to students with educator-identified needs (e.g., glossaries in 13 languages) to support specific kinds of engagement; and
- Accommodations for individual students with documented IEP or 504 plans (e.g., braille versions).

These supports are designed to ensure that each student engaging with the assessment has the best opportunity to show what they know and can do. The intention is to *enhance* construct comparability, rather than detract from it by ensuring that differences in performance reflect differences in skill, not access. While these supports introduce some variability in test administration procedures, ranging from minor (e.g., notepad use) to more substantial (e.g., braille), they increase fairness by enabling students to more accurately demonstrate competency. This approach illustrates how standardized assessment systems can integrate flexibility to better serve a diverse population without compromising measurement integrity.

03

Theme Three: Expanding Construct Definitions to Account for Social, Cultural, and Linguistic Contributions to Learning and Performance

Across arguments for increasing the relevance and personalization of assessments for specific learners, there has been an implicit—and at times explicit—call to rethink both what assessments are measuring and how those measurements are interpreted. This rethinking includes calls to:

1. Expand the range of assessment constructs to better reflect how learning and development actually occur;
2. Incorporate culturally and community-specific priorities and goals for student learning; and
3. Develop more trustworthy and inclusive measures of what students know and can do, even within narrowly defined domains.

Together, these shifts aim to ensure that assessments more accurately reflect the full breadth of student learning and experience.

Lee (2025) contends that the Science of Learning and Development (SoLD)—which synthesizes interdisciplinary research from human development, psychology, neuroscience, and the learning sciences to offer a comprehensive understanding of the diverse factors that shape learning across the lifespan—requires a fundamental rethinking of the aims and design of educational assessments. That rethinking has the potential to yield more actionable, equitable, and ecologically valid insights. Lee argues that persistent disparities in assessment outcomes—by race/ethnicity, gender, and socioeconomic status—stem not only from inequitable opportunities to learn but also from the limitations of existing assessment systems. Current summative, interim, and formative approaches fail to capture the full range of influences on how people learn and demonstrate knowledge. These influences include cultural identity, perceptions of the task and setting, emotional salience, epistemological beliefs, mindsets, and self-efficacy.



Moreover, Lee critiques the dominance of narrow, Eurocentric definitions of disciplinary knowledge in U.S. education, which constrain both teaching and assessment. She argues that expanding our conceptions of what counts as knowledge—in ways that are more culturally and contextually responsive—could reduce disparities, foster a more holistic view of learners' capabilities, and better recognize the strengths and knowledge systems students from historically marginalized communities bring to school.

Lee's challenge to narrow definitions of disciplinary knowledge in current large-scale assessment systems is echoed in many other chapters in the volume. For example, Welch and Dunbar (2025) examine opportunities for integrating SCRA into federally mandated state assessments. Based on their analysis, they argue that current interpretations of alignment in both item development and item/test evaluation (e.g., for purposes of federal peer review) privilege overly narrow conceptions of what it looks like to demonstrate performance relative to established standards.

Kūkea Shultz and Englert (2025) describe how ideas related to reimagining what is assessed on large-scale, federally mandated assessments are factored into a culturally-specific operational assessment: the Kaiapuni Assessment of Educational Outcomes (KĀ'EO), the accountability assessment for students attending Hawaiian language immersion programs. Developing KĀ'EO required explicitly countering the standard, monocultural worldviews that govern most large-scale assessment development processes because central to KĀ'EO's purpose is the reclamation of Hawaiian culture and language; this reclamation required defining constructs in ways that are responsive to the local community's language, culture, and information needs while still meeting federal requirements for showing progress in mathematics, language arts, and science.

KĀ'EO also relies upon a deeply collaborative and relational approach, placing educators and community members within the Kaiapuni system as decision-makers throughout the assessment development and validation process. Kūkea Shultz and Englert (2025) posit community validity—reflective of the intentional processes governing KĀ'EO's creation and validation—as a framework for assessment development that positions assessment as an activity that serves communities as a first-order principle.

Many scholars (e.g., Lee 2025, Nelson-Barber and Trumbull 2025, Kūkea Shultz and Englert 2025) emphasize the need to disrupt the dominance of White, Eurocentric cultural norms that shape not only how assessments are designed, but also how students experience schooling more broadly.

Across testing contexts, many of the book's chapter authors contend that transforming assessment requires a fundamental shift in what we value, how we define competence, and whose knowledge is recognized. By bridging definitions of what we measure—and what we value—with the full diversity of learner experiences, they suggest that assessment systems can become more inclusive, relevant, and empowering.

04

Theme Four: Assessment Development Processes that Account for Sociocultural Goals

A fourth theme that runs through the book is the need for and use of frameworks and tools to guide assessment development and score interpretation in ways that align with the conceptual and evidence-based recommendations for SCRA. Frameworks have long been used with standardized tests for such purposes. Classical test theory (Gulliksen, 1950), item response theory (IRT; Lord, 1980), evidence-centered design (ECD; Mislevy et al., 2003), and the argument-based approach to validation (Kane, 1992) are widely used examples.

Frameworks are very useful for purposes of SCRA because SCRA presents no shortage of problems, including ones related to assessment design, development, analysis, and interpretation. Key to SCRA design is taking account of examinee sociocultural characteristics to allow individuals to demonstrate better what they know and can do. Sato (2025) offers a framework for design that focuses on deeper levels of culture, with the intention of accounting for those factors (e.g., values, norms, beliefs, language, social structure/dynamics, milieu) that affect students' meaning making and their representations of knowledge. The chapter presents a *sociocultural dimensions matrix* describing personal orientations. The matrix should be of use in designing more inclusive measurement targets, tasks, and scoring rules, as well as for guiding interpretations of diverse student performance.

In her chapter, Ebe (2025) continues the concern with more effectively accounting for sociocultural orientation in assessment. In line with Sato's (2025) chapter, as well as with much reading research (Lee, 2025; Wang et al., 2025), Ebe's premise is that comprehension depends critically upon relevant prior knowledge. We more quickly and completely understand text that draws upon what we already know. Moreover, what we know is culturally shaped. Thus, knowing the cultural relevance of text should help in SCRA design and score interpretation. Ebe's Cultural Relevance Rubric offers a framework for such activities.

A general framework for guiding the development of SCRAs is offered in the chapter written by Badrinarayan and Darling-Hammond (2025). The framework was derived from a review of state and national attempts to account for sociocultural factors in large-scale assessment systems. Badrinarayan and Darling-Hammond's framework consists of five features intended to be used as "... a heuristic for defining the potential landscape of culturally conscious assessment systems at scale" (p. 362).



Sato's (2025), Ebe's (2025), and Badrinarayan and Darling-Hammond's (2025) chapters principally focus on frameworks and tools for SCRA development. In contrast, Moses (2025) deals with issues of analysis and interpretation. He compares sociocultural theories with large-scale assessment practices for test scoring and linking. He notes that tests developed from a socioculturally responsive perspective may vary across examinees, presenting different questions or posing them in ways that fit test takers' backgrounds. Large-scale assessment practices, in contrast, ideally require design to the same constructs and specifications to maintain consistent score meaning. Moses depicts the challenge as a tradeoff between comparability across examinees vs. validity for a particular use and/or group. He suggests ways this local validity vs. broad comparability tension might be resolved.

Like Moses, Mislevy et al. (2025) deal with the issues posed by SCRA for analysis and score interpretation, but additionally with the implications of SCRA for development. These authors propose a lens that connects the logical assessment-argument structures of Evidence-Centered Design (Mislevy et al., 2003) with the sociocultural aspects of tasks (e.g., aspects that are construct essential, related to sociocultural background, ancillary, enabling or restricting, genre specific). That combination allows for analyzing the relations among tasks, students, purposes, and inferences to facilitate assessment design decisions.

Conclusion and Next Steps for SCRA

Across the many contributions to the source volume *Socioculturally Responsive Assessment: Implications for Theory, Measurement, and Systems-Level Policy* (Bennett, Darling-Hammond, & Badrinarayan, 2025b), one message emerges with clarity: SCRA requires more than technical improvements or inclusive messaging—it demands a transformational shift in how we conceptualize, design, implement, and interpret educational assessments at scale.

This shift begins by re-centering the foundational assumption that learning is inseparable from culture, identity, and context, and that assessment must reflect this reality. The chapters in the volume demonstrate how assessment systems grounded in sociocultural responsiveness can be designed to recognize multiple ways of knowing, support meaningful engagement with academic content, and honor the full range of students' cultural and linguistic repertoires. Whether by increasing the relevance of content, personalizing assessment processes, or broadening construct definitions, the work described challenges dominant paradigms that have historically marginalized non-dominant learners.

Importantly, contributors show that technical quality and sociocultural responsiveness are not mutually exclusive. Through examples such as the KĀ'EO assessment in Hawai'i, and Smarter Balanced's universal design features, the chapters illustrate how construct comparability can be preserved or even strengthened when design decisions are guided by community-informed theories of action, clearly articulated validity arguments, and inclusive development practices. This consideration includes not only attention to item and task development, but also to scoring practices, data use, and communication of results—all of which must be revisited to better serve students and communities.

To that end, the volume also highlights the value of new frameworks and analytical tools—including cultural relevance rubrics, sociocultural design matrices, and approaches to UNDERSTANDARDIZATION—that help operationalize SCRA in practical terms. These tools underscore the need to decenter monolithic conceptions of standardization and instead embrace models that prioritize construct validity over procedural uniformity, allowing for more nuanced, community-centered, and asset-based approaches to large-scale measurement.

While the field continues to wrestle with trade-offs between comparability and responsiveness, efficiency and authenticity, and alignment to standards and cultural relevance, the authors in this volume make clear that we must move beyond binary thinking. As Moses, Sato, Kūkea Shultz and Englert, and others argue, validity arguments must be both technically sound and socially credible, grounded not just in statistical evidence, but in the lived realities and values of the communities assessments are intended to serve.

Whereas this chapter summary highlights common themes across a wide range of ideas in the book, there are many open questions with regard to SCRA. Some of these questions include:

- **Whose culture should be centered and valued in assessment design?** This foundational question arises across scholarly, practical, and policy discussions surrounding SCRA. While some scholars argue that SCRA should create fairer assessments for all learners by expanding inclusivity (e.g., Lee, 2025; Zandvakili & Gordon, 2025), others advocate for explicitly centering the needs and experiences of historically marginalized student groups as a primary goal (e.g., Randall et al., 2022, Randall 2023). This tension becomes more complex as SCRA efforts increasingly seek to account for the nuanced, intersectional identities of learners. Advancing this work will require deeper inquiry into how cultural representation decisions are made within assessment systems—who is involved, what criteria are used, and how trade-offs are weighed. Such efforts are critical not only for guiding the design and implementation of SCRA-aligned assessments but also for informing broader policy decisions about inclusion, accountability, and equity in education.
- **How do different use cases for large-scale assessment govern appropriate trade-offs for incorporating SCRA into assessment design?** There are many different reasons for including SCRA in assessment design, ranging from humanizing students' experiences with assessment to generating more trustworthy scores. When these reasons are further contextualized by the uses of assessments—both intended and actual—principled decisions about how SCRA is incorporated into different assessment designs can be more effectively made. These considerations can also help identify which aspects of assessment design—such as item content, administration procedures, or reporting formats—require the most focused attention to ensure equitable and meaningful use.
- **How can emerging technologies support SCRA?** As Bennett et al. (2025a) describe, there are significant opportunities for emerging technologies, including artificial intelligence, to better support SCRA. For example, generative AI might support real-time personalization; more efficient generation of item pools that reflect greater linguistic, cultural, topical, and social diversity; the creation of immersive simulations that better capture social reasoning and authentic engagement to support more valid assessments of complex, deeper learning competencies; and the synthesis of evidence across multiple demonstrations of learning to produce more holistic representations of student capabilities. Additionally, AI-powered technologies show promise in developing flexible scoring mechanisms capable of interpreting responses expressed through varied modalities—such as spoken language, prose, bullet points, graphics, or symbolic representations. These technologies also may help in producing responsive and interactive reporting systems that are better attuned to users' linguistic preferences, implementation settings, and immediate information needs.

Ultimately, the volume underscores that transforming large-scale assessment is both necessary and possible, and is already underway. It is a call to action for assessment developers, policy leaders, educators, and researchers to build systems that reflect a pluralistic vision of learning, one in which all students are seen, heard, and empowered. The future of large-scale assessment must not only measure what students know and can do, but also support who they are and who they aspire to become. If we are to create assessment systems that are truly equitable, valid, and educationally valuable, the work of SCRA cannot be peripheral—it must be central to our reimagining of assessment systems.

References

- Badrinarayan, A., & Darling-Hammond, L. (2025). Enacting socioculturally responsive assessment design at scale: Approaches and policies for large-scale systems. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 336–366). Routledge.
- Badrinarayan, A., Lyons, S., Miranda, A., & Klyachkina, A. (2025). Leveraging students' cultural and linguistic assets for assessment: A framework for culturally conscious assessment in the Chicago Public Schools. *Educational Assessment, 30*(3), 1–26. <https://doi.org/10.1080/10627197.2025.2497775>
- Bennett, R. E. (2023). Toward a theory of socioculturally responsive assessment. *Educational Assessment, 28*(2), 88–104. <http://dx.doi.org/10.1080/10627197.2023.2202312>
- Bennett, R. E. (2024). Personalizing assessment: Dream or nightmare? *Educational Measurement: Issues and Practice, 43*(4), 119–125. <https://doi.org/10.1111/emip.12652>
- Bennett, R. E. (2025). A descriptive review of culturally responsive, socioculturally responsive, and related assessment conceptions. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 11–28). Routledge.
- Bennett, R. E., Darling-Hammond, L., & Badrinarayan, A. (2025a). *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy*. Routledge.
- Bennett, R. E., Darling-Hammond, L., & Badrinarayan, A. (2025b). Socioculturally responsive assessment: Present and future. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 434–442). Routledge.
- Ebe, A. (2025). Examining the relationship between the cultural relevance of text and reading proficiency: Using a cultural relevance rubric in reading assessment. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 168–179). Routledge.
- Escoffery, D. S., Fletcher, K. E., & Stone-Danahy, R. A. (2025). "A search for my voice": Socioculturally responsive assessment in AP Art and Design. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 262–282). Routledge.
- Gulliksen, H. (1950). *Theory of mental tests*. Wiley.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112*, 527–535. <https://doi.org/10.1037/0033-2909.112.3.527>
- Kūkea Shultz, P., & Englert, K. (2025). A path to transforming the assessment landscape. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 306–335). Routledge.
- Lee, C. D. (2025). Implications of the science of learning and development (SoLD) for assessments in education. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 29–49). Routledge.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Lawrence Erlbaum.
- Michel, R., & Shyyan, V. (2025). Accessibility as a core value for locally responsive assessments. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 247–261). Routledge.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *A brief introduction to evidence-centered design* (Research Report No. RR-03–16). Educational Testing Service.
- Mislevy, R. J., Oliveri, M. E., Slomp, D., Crop Eared Wolf, A., & Elliot, N. (2025). 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. 199–242). Routledge.

References

- Moses, T. (2025). Linking responsive assessments: Challenges and possibilities. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 180–198). Routledge.
- Nasir, N. S., Lee, C. D., Pea, R. D., & McKinney de Royston, M. (Eds.). (2020). *Handbook of the cultural foundations of learning*. Routledge.
- National Academies of Sciences, Engineering, and Medicine. (2018). *How People Learn II: Learners, Contexts, and Cultures*. The National Academies Press. <https://doi.org/10.17226/24783>.
- Nelson-Barber, S., & Trumbull, E. (2025). Assessment for social justice with a focus on American Indigenous students. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 76–98). Routledge.
- Raji, M. O., & Baidoo-Anu, D. (2025). Socioculturally responsive post-secondary entrance examination: Implications for equitable assessment design in sub-Saharan Africa. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 399–414). Routledge.
- Randall, J. (2023). It Ain't Near 'Bout Fair: Re-envisioning the bias and sensitivity review process from a justice-oriented, antiracist perspective. *Educational Assessment*. Advance online publication. <https://doi.org/10.1080/10627197.2023.2223924>
- Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting White supremacy in assessment: Toward a justice-oriented, antiracist validity framework. *Educational Assessment*, 27(2), 170–178. <https://doi.org/10.1080/10627197.2022.2042682>
- Sato, E. (2025). Born socioculturally responsive assessment: An approach to design and development. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 115–143). Routledge.
- Sireci, S. G., Crespo Cruz, E., Suárez-Álvarez, J., & Rodríguez Matos, G. (2025). Understanding UNDERSTANDardization research. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 415–433). Routledge.
- Skerrett, A. S., Pacheco, M., Hinchman, K. A., Cervetti, G. N., Pearson, P. D., & Greenleaf, C. (2025). Socioculturally responsive large-scale assessment: The case of the 2026 NAEP reading framework. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 283–305). Routledge.
- Solano-Flores, G., & Ruiz-Primo, M. A. (2025). Cultural validity: Conceptual and methodological considerations 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. 50–75). Routledge.
- Wang, Z., Sparks, J. R., Walker, M., O'Reilly, T., & Bruce, K. (2025). Group differences across scenario-based reading assessments: Examining the effects of culturally relevant test content. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 369–398). Routledge.
- Welch, C., & Dunbar, S. (2025). Designing socioculturally responsive assessments within the mandated bookends of K–12 assessment systems. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 144–167). Routledge.
- Zandvakili, E., & Gordon, E. W. (2025). Transforming educational assessment: Implementing equitable and inclusive strategies. In R. E. Bennett, L. Darling-Hammond, & A. Badrinarayan (Eds.), *Socioculturally responsive assessment: Implications for theory, measurement, and systems-level policy* (pp. 99–112). Routledge.

About the authors

Aneesha Badrinarayan is the Principal Consultant at ABG Consulting, where she partners with state and district leaders, assessment developers, and policymakers to reimagine how teaching, learning, and assessment work together. A behavioral neuroscientist by training, she brings decades of expertise spanning assessment design, STEM education, policy, and product development — with a consistent focus on building systems that are instructionally relevant, equitable, and future-oriented. Before founding ABG Consulting, Aneesha led assessment innovation initiatives at Education First, the Learning Policy Institute, and Achieve, where she led initiatives across 15 states, helped shape the 2028 NAEP Science Framework, guided federal policy on learning-first assessment, and pioneered work at the intersection of equitable assessment, accountability design, and artificial intelligence. Aneesha holds degrees from Cornell University and the University of Michigan. She is the author of *Surfacing Brilliance Through Meaningful Science Assessment* (2025).

Randy E. Bennett is Principal with Assessment Innovation Matters. His recent work centers on personalized assessments and, relatedly, assessments that are “born socioculturally responsive.” From 1999–2005 he directed the National Assessment of Educational Progress (NAEP) Technology-Based Assessment project, which included the first administration of computer-based performance assessments to nationally representative samples of U.S. school students and the first use of logfile data in such samples to measure problem-solving processes. From 2007–2016, he directed the CBAL research initiative (Cognitively Based Assessment of, for, and as Learning), which created theory-based summative and formative assessment to model good teaching and learning practice. He is a past president of the International Association for Educational Assessment and of the National Council on Measurement in Education (NCME). He is a fellow of the American Educational Research Association (AERA) and an elected member of the National Academy of Education, as well as recipient of the NCME Bradley Hanson Contributions to Educational Measurement Award, the Teachers College Columbia University Distinguished Alumni Award, the AERA E. F. Lindquist Award, and the AERA Cognition and Assessment SIG Award for Outstanding Contribution to Research in Cognition and Assessment.

Linda Darling-Hammond is the Charles E. Ducommun Professor of Education, Emeritus, at Stanford University and founding president of the Learning Policy Institute, where she leads research and policy initiatives focused on educational equity, teacher quality, and effective school reform. A nationally renowned scholar, she has authored more than 30 books and hundreds of publications on teaching, learning, and education policy. Darling-Hammond's career has centered on advancing evidence-based policies that improve access to high-quality learning opportunities for all students. She has served as president of the CA State Board of Education since 2019. In that role, she has guided the state's efforts to strengthen curriculum, assessments, and teacher preparation. Earlier, she directed the Stanford Center for Opportunity Policy in Education and the National Commission on Teaching and America's Future, influencing reforms in teacher development and accountability systems across the U.S. Recognized as one of the most influential voices in education, she has advised federal and state leaders on issues ranging from school funding to equitable assessment design. Darling-Hammond continues to champion the creation of schools that support deep learning, social-emotional growth, and equitable outcomes for every child.

About the Study Group

The Study Group exists to advance the best of artificial intelligence, assessment, and data practice, technology, and policy; uncover future design needs and opportunities for educational systems; and generate recommendations to better meet the needs of students, families, and educators.

Date of Publication

March 2026

Images

All imagery is public-domain content provided by the Study Group.

Citation

Badrinarayan, A., Bennett, R. E., & Darling-Hammond, L. (2026). *Perspectives on socioculturally responsive assessment in large-scale systems: A chapter summary*. The Study Group.

Licensing

This case study is [based on a chapter](#) that has been made available under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International ([CC BY-NC-ND](#)) license. All photos in this chapter were provided by the Study Group from public domain sources.