Handout 1: Instruments for Measuring the Cultural Responsiveness of Students' Educational Experiences

Culturally responsive education (CRE)[1] "advances equity and social justice by: centering and valuing students' cultures and identities; using rigorous and culturally relevant curriculum and anti-oppressive teaching practices; building strong, positive relationships between students, families, and school staff; and, supporting students to develop the knowledge, skills, and vision to transform the world toward liberation."[2] In 2020, members of REL Appalachia's Improving Postsecondary Transitions partnership expressed interest in checklists or tools for assessing and enhancing the cultural responsiveness of students' educational experiences. In response, REL Appalachia researchers scanned the literature to identify five research-based instruments educators can use to assess the cultural responsiveness of their students' educational experiences.

Table 1 describes each of the instruments, including the dimensions(s) of CRE each instrument assesses, the type of research the authors cite as informing its development, and a summary of its reliability and validity. CRE demands a comprehensive understanding of the educational system and how each facet can affect the experiences of students of color, which is why it is important to consider these interconnected dimensions in selecting and using the resources.

Box 1 provides definitions for the spectrum of research types that contribute to the development and testing of interventions and strategies that the authors cite as informing each instrument's development. To date, research that explores the causal impact of CRE on student outcomes is limited.[3,4] Much of the research consists of foundational research studies that describe CRE without establishing a link to student outcomes, and small-scale, early-stage, or exploratory research showing correlational relationships between CRE and student achievement, persistence, or affective outcomes such as motivation or confidence.[5] In furthering the evidence-based research connecting CRE to student outcomes, CRE scholars argue that outcomes should not be limited to standardized test scores, if standardized test scores should be included as an outcome at all, given cultural bias concerns in standardized testing.[3,5,6]

Box 2 provides definitions for "reliability" and "validity" used to assess each instrument's strength in measuring the cultural responsiveness of students' educational experiences. It is

---

[1] CRE is inclusive of the closely related terms culturally responsive teaching and culturally relevant pedagogy.
[2] Culturally Responsive Education Hub. (n.d.) https://crehub.org/
[3] Hill, H. (2020, March 6). Culturally responsive teaching is promising. But there's a pressing need for more research. *Education Week.* https://www.edweek.org/leadership/opinion-culturally-responsive-teaching-is-promising-but-theres-a-pressing-need-for-more-research/2020/03
[4] Sleeter, C. (2012). Confronting the marginalization of culturally responsive pedagogy. *Urban Education, 47,* 562–584. doi:10.1177/0042085911431472
[5] Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. *Review of Educational Research, 86(1)*, 163-206.
[6] Kim, K. H., & Zabelina, D. (2015). Cultural bias in assessment: Can creativity assessment help? *International Journal of Critical Pedagogy, 6(2),* 129–147.

important to consider reliability and validity when selecting an instrument. Although authors should ideally collect and report on evidence of an instrument's reliability and validity, not all do. Often, extensive research is required to generate reliability and validity evidence. Instruments that do not report reliability and validity currently may be updated in the future, once instruments are tested in different settings and with different populations of educators and students.

REL Appalachia researchers relied on author reports of reliability and validity, where available. In some cases, REL Appalachia researchers independently inferred evidence of content validity when sufficient detail was provided in the instrument documentation. Specifically, REL Appalachia researchers examined whether authors describe each aspect of the construct they are attempting to measure with supporting citations to the theoretical or foundational research supporting these ideas, differentiate the construct of interest from other constructs, and/or demonstrate how the items contained in the instrument align to each aspect of the construct.

Table 1. Instruments for measuring the cultural responsiveness of students' educational experiences.

Instrument/Tool	Author	Excerpted Description	Intended User/Audience	Dimension(s) of CRE Addressed[7]	Relevant/ Supporting Literature Cited in Instrument Description[8]	Evidence of Reliability and Validity of Instrument[9]
"And they cared": How to Create Better, Safer Learning Environments for Girls of Color	The Education Trust and National Women's Law Center (2020)	This guide aims to provide decisionmakers with a common language and practices that can be used to reform exclusionary discipline policies and improve school climate to help address the needs of girls of color—especially Black girls. It includes a checklist, so decisionmakers can assess what changes they can make to create positive school climates at the school, district, or state level.	Self-reflection for school, district, and state leadership	Administration and policies	Yes: Foundational Research; Early-Stage or Exploratory Research	No reliability or validity information supporting the development of the checklist was provided in the available documentation.
Assessing Bias in Standards and Curricular Materials	Midwest and Plains Equity Assistance Center (2017)	This tool enables users to determine the extent to which developed standards and curricular materials reflect educational equity. The tool provides guidance in reviewing standards and curricular materials using equity-oriented domains. It also includes a scoring and analysis guide to assist with the evaluation process.	Self-report tool for teachers, school administrators	Instructional materials and assessments Teaching practices and activities	Yes: Foundational Research	No reliability evidence supporting the development of the tool was provided in the available documentation. Content validity: adaptation of an existing tool[10]

---

[7] Adapted from Richards, H. V., Brown, A. F., Forde, T. B. (2007). Addressing diversity in schools: Culturally responsive pedagogy. *Teaching Exceptional Children, 39(3)*, 64-68..

[8] See Box 1 for a listing of research types and definitions.

[9] See Box 2 for a listing of reliability and validity terms and definitions.

[10] The Assessing Bias in Standards & Curricular Materials Tool is adapted from the 2016 Equity Leaders Institute (ELI) by The Great Lakes Equity Center and David Sadker's (n.d.) *Some Practical Ideas for Confronting Curricular Bias.*

Instrument/Tool	Author	Excerpted Description	Intended User/Audience	Dimension(s) of CRE Addressed[7]	Relevant/ Supporting Literature Cited in Instrument Description[8]	Evidence of Reliability and Validity of Instrument[9]
[Culturally Responsive Curriculum Scorecard](#)	NYU Metro Center (2019)	NYU Metro Center designed this tool to help parents, teachers, students, and community members determine the extent to which their schools' English Language Arts curricula are (or are not) culturally responsive. The scorecard comes with an accompanying toolkit which provides resources and guidance for students, parents, communities, educators, and administrators on next steps to make your classrooms and schools more culturally responsive, after you have used the Culturally Responsive Curriculum Scorecard.	Self-report tool for parents, teachers, students, and community members	Instructional materials and assessments	Yes: Foundational Research; Early-Stage or Exploratory Research	No reliability evidence supporting the development of the scorecard was provided in the available documentation. Content validity: development of the scorecard was based on existing resources.[11]
[Culturally Responsive Teaching: A Reflection Guide](#)	New America (2020)	This resource is intended to support the reflective practice and ongoing learning of culturally responsive teachers. It can help teachers assess their personal strengths and develop a plan to sharpen their practice. Additionally, this guide can and should be used by those who support teachers.	Self-reflection for teachers	Teaching practices and activities Teacher attitudes and beliefs	Yes: Foundational Research; Early-Stage or Exploratory Research	No reliability evidence supporting the development of the competencies or reflection questions was provided in the available documentation. Content validity: each competency is described; and associated reflection questions are listed.

[11] To create the Culturally Responsive Curriculum Scorecard, the authors "drew upon a wide variety of existing resources, including multicultural rubrics, anti-bias rubrics, textbook rubrics, and rubrics aimed at creating cultural standards for educators, determining bias in children books and examining lesson plans (ADEED, 2012; Aguilar-Valdez, 2015; Grant & Sleeter, 2003; Lindsey et al, 2008; NCCRES, 2006; Rudman, 1984; World View, 2013)" (NYU Metro Center, 2019, p. 4).

Instrument/Tool	Author	Excerpted Description	Intended User/Audience	Dimension(s) of CRE Addressed[7]	Relevant/ Supporting Literature Cited in Instrument Description[8]	Evidence of Reliability and Validity of Instrument[9]
Culture in the Classroom: Standards, Indicators, and Evidences for Evaluating Culturally Responsive Teaching	Southeast Regional Resource Center (SERRC): Alaska's Educational Resource Center (2015)	This document was developed as part of SERRC's Project CREATE to develop and support a teacher evaluation framework aligned with research-based instructional models such as Marzano's Art & Science of Teaching, and Danielson's Framework for Teaching. School districts and personnel will find this publication useful in assisting educators in the teacher evaluation process. It provides examples of what culturally responsive instruction and practice could look like. Teachers will also find it helpful in planning for meaningful, effective lessons, activities, and strategies that meet the cultural standards for educators.	Self-reflection for teachers and for evaluation and feedback by teacher evaluators	Teaching practices and activities	Yes: Foundational Research	No reliability evidence supporting the development of the evaluation framework was provided in the available documentation. Content validity: cultural descriptions, indicators, and evidence statements aligned to the existing Alaska Cultural Standards for Educators were developed by a large multidisciplinary team.

Box 1. Types of Research

Foundational Research: Provides the fundamental knowledge that may contribute to improved learning and other relevant education outcomes. Studies of this type seek to test, develop, or refine theories of teaching or learning and may develop innovations in methodologies and/or technologies that will influence and inform research and development in different contexts. Foundational Research studies may examine phenomena without establishing an explicit link to education outcomes.

Early-Stage or Exploratory Research: Examines relationships among important constructs in education and learning to establish logical connections that may form the basis for future interventions or strategies to improve education outcomes. These connections are usually correlational rather than causal.

Design and Development Research: Develops solutions to achieve a goal related to education or learning, such as improving student engagement or mastery of a set of skills. Research projects of this type draw on existing theory and evidence to design and iteratively develop interventions or strategies, including testing individual components to provide feedback in the development process.

Efficacy Research: Allows for testing of a strategy or intervention under "ideal" circumstances, including with a higher level of support or developer involvement than would be the case under normal circumstances. Efficacy Research studies may choose to limit the investigation to a single population of interest.

Effectiveness Research: Examines effectiveness of a strategy or intervention under circumstances that would typically prevail in the target context. The importance of "typical" circumstances means that there should not be more substantial developer support than in normal implementation, and there should not be substantial developer involvement in the evaluation of the strategy or intervention.

Scale-Up Research: Examines effectiveness in a wide range of populations, contexts, and circumstances, without substantial developer involvement in implementation or evaluation. As with Effectiveness Research, Scale-up Research should be carried out with no more developer involvement than what would be expected under typical implementation.

Source: Institute of Education Sciences (IES) and National Science Foundation (NSF). (2013). *Common guidelines for education research and development.* Authors. https://ies.ed.gov/pdf/CommonGuidelines.pdf

Box 2. Reliability and Validity Terms and Definitions

Reliability: Whether an instrument consistently measures the experience across respondents, time, or raters. Information that could support reliability includes several families of statistics, including measures of internal consistency such as Cronbach's α, omega or test-retest associations, and inter-rater reliability. **The conventionally accepted criterion of reliability for Cronbach's α is ≥ .70.** Reliability statistics, such as Cronbach's α, are used to examine the likelihood of an instrument generating similar scores under consistent conditions.

Validity. Whether an instrument measures what it is intended to measure and whether the inferences drawn from an instrument are appropriate.

    Content validity. Whether items for an instrument contain content that adequately describes the experience(s).

    Structural validity. Whether an instrument's items relate statistically to the experience being measured.

    External validity. Whether there are correlations between scores from the instrument and scores from other instruments measuring similar experiences.

    Consequential validity. Whether scores generated from an instrument are associated with the intended consequences of using the instrument, such as improving student outcomes.

    Generalizability validity. Whether scores from the instrument are correlated with other modes of measurement for the same experience, such as self-reported or observational information. Unlike external validity, which explores correlations between scores from different instruments measuring similar experiences, generalizability explores correlations between scores from different modes of measurement of the same experience.

Adapted from: Cox, J., Foster, B., & Bamat, D. (2019). A review of instruments for measuring social and emotional learning skills among secondary school students (REL 2020–010). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. http://ies.ed.gov/ncee/edlabs