Designing Culturally Responsive Online Assessments for Equity-Deserving Students

Kim Koh
University of Calgary
khkoh@ucalgary.ca

Abstract

In this article, the author offers a brief review of the theoretical foundations, historical origins, and development of culturally responsive assessment (CRA). The concept of authentic performance assessment is discussed with its crucial role in serving as a form of CRA to support students’ learning in online learning environments. Drawing on the fields of measurement, assessment, and technology, the author highlights some key considerations in the design and implementation of culturally responsive online assessments for equity-deserving students.

Keywords: Culturally responsive assessment, online learning environments, assessment design, equity-deserving students, measurement, technology

The COVID-19 pandemic brought about a wide range of social, economic, and political challenges globally, and the education sector was severely impacted as a result. At an individual level, many students experienced significant learning setbacks, increased absenteeism, and mental health challenges due to prolonged school closures and social isolation. Additionally, PreK-12 teachers and higher education instructors faced challenges in swiftly adapting their lessons and assessments for online delivery. Some teachers reportedly experienced high levels of technostress due to the many responsibilities they were expected to fulfill during the pandemic (Zheng et al., 2022).

At the onset of the pandemic, strict regulations on physical distancing forced schools and universities worldwide to close. To mitigate learning loss, PreK-12 schools and higher education institutions implemented online instruction. School teachers and university instructors were given limited time to modify their instructional plans for online delivery. Similarly, classroom assessments and end-of-term exams had to be redesigned for online completion. Both teachers and students had to rapidly acquire technological skills, with students being compelled to complete assignments and take exams online. For many teachers and students, they had to pick up both assessment and technological skills within a short amount of time. The sudden shift to online instruction and assessment disadvantaged students in rural areas or with limited access to computers and the internet. Many of these students, often from low-income, working class or immigrant families, had already faced marginalization. More specifically, such students included English language learners (ELLs), students of color, Indigenous students, first-generation university students, girls, and women. In the United States and Canada, they are also referred to as equity-deserving students who are historically underrepresented and underserved.

Even before the pandemic, the issues of assessments and equity had already gained significant attention in educational research and policymaking in countries such as the United...
States, Canada, the United Kingdom, Australia, and New Zealand, which have a sizable population of Indigenous students and an influx of immigrant students. Over the past decade, Canada has assisted a record number of refugees from Iraq, Syria, and the Ukraine to resettle within its borders. Public school PreK-12 classrooms in Canada have become more linguistically and culturally diverse because of governmental policies accepting immigrants and refugees from war-torn countries like the Ukraine. According to Statistics Canada (2022), nearly one in four individuals (23.0%) counted during the 2021 Census are immigrants. In Canada, Africans represent the largest group of immigrants. In the United States, immigrant children under the age of 18 constitutes 26% of the PreK-12 student population, and one in four children have at least one immigrant parent and speak a language other than English at home (Pew Research Center, 2015). According to the U.S. National Center for Education Statistics (2023), the percentage of public-school students who were ELLs increased from 9.2% (4.5 million students) in fall 2010 to 10.3% (5.0 million students) in fall 2020.

These statistics highlight the presence of a substantial group of youth and school-age children who are newcomers and/or ELLs and must integrate to the dominant culture in order to settle effectively in their new home and school. Considering that the academic qualifications and professional experiences of most refugees or immigrant parents are not recognized in the local labor market, they typically have lower incomes and lack resources to support their children. Indigenous and racially minoritized students are also historically marginalized. They face greater economic, social, cultural, and political barriers, as well as educational disparities, as compared to most white, middle-class students.

In early May 2023, the World Health Organization declared the end of the pandemic. At the time of writing this article, schools and higher education institutions worldwide have resumed in-person instruction. Some may have adopted a blended learning approach or a HyFlex model, providing students with the option to choose between online or in-person learning (Lockee, 2021).

In a post-pandemic era, assessing equity-deserving students’ learning in an online or blended environment poses even greater challenges and complexities compared to assessment in traditional physical, face-to-face classrooms. Various issues have become increasingly important to consider, including online learning, hybrid learning, racial injustice, decolonizing education, marginalized communities, social inequality, diversity and inclusion, and accessibility. The author of this article asserts that assessments that aimed at supporting the learning of all students must encompass each of these aspects in their design, adaptation, and implementation. To ensure the validity of inferences made about students’ learning, the measurement of their progress and outcomes must be responsive and culturally relevant. This entails considering fairness and equity issues even during the design of assessment tasks or test items. The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014) provide specific guidelines for ensuring the technical rigor of assessments. These considerations are especially important in online learning environments, given the increased number and diversity of learners, as well as the features and limitations of new technologies such as virtual reality, augmented reality, and artificial intelligence (Koh, Chapman & Lam, 2022).
Additionally, student learning outcomes encompass multiple dimensions that extend beyond disciplinary knowledge. For example, assessing STEM (science, technology, engineering, and mathematics) competencies requires assessing students’ ability to apply integrated STEM subjects to solve complex, real-world problems, thus emphasizing interdisciplinary or transdisciplinary problem-solving in context (Koh, Chapman & Liu, 2022). The necessity to assess students’ cross-cutting competencies or 21st century skills, including critical thinking, creative problem-solving, collaboration, communication, and civic and social responsibility, suggests that the assessment of student learning should encompass both the cognitive and affective domains of learning. This includes: considerations of students’ social and emotional well-being; the ability to self-regulate their learning; their self-efficacy or confidence, motivation, resilience, and growth mindset. Consequently, designing and implementing classroom assessment tasks has become more challenging for PreK-12 teachers and higher education instructors.

In this article, the author offers a brief review of the theoretical foundations, historical origins, and development of culturally responsive assessment (CRA). It will focus on how authentic performance assessments can serve as a form of CRA to support students’ learning in online learning environments. Furthermore, the article will highlight some key considerations in the design and implementation of culturally responsive online assessments for equity-deserving students.

Historical Roots and Evolution of Culturally Responsive Assessment

A literature review was conducted using the Academic Search Complete and Google Scholar databases to explore the conceptual definitions, theoretical foundations, and practical applications of CRA. There is a scant body of empirical literature on the design and implementation of CRA in elementary schools. However, some studies examined CRA in relation to underrepresented students in high schools (e.g., Bahar & Maker, 2020; Miranda & Cherng, 2018) and higher education institutions, including teacher education (e.g., Henning & Lundquist, 2022; Montenegro & Jankowski, 2017; Walker et al., 2021). It is worth noting that CRA is associated with culturally responsive pedagogy and culturally responsive evaluation (CRE). According to Hood et al. (2015),

The early roots of CRE began in education, specifically in the work of Carol Lee (1990) and Gloria Ladson-Billings (1995a,b) on culturally responsive pedagogy in conjunction with the work of Edmund Gordon (1998) and Sylvia Johnson (1998) in educational assessment. Hood (1998a) extended this thinking from culturally responsive pedagogy to culturally responsive assessment, and subsequently to culturally responsive evaluation. (p. 285)

Hood et al. (2015) further highlight that “the bridge between culturally responsive assessment to culturally responsive evaluation was built within validity theory” (p. 285). Hood’s conceptualization of CRA has been influenced by Samuel Messick’s (1989) definition of construct validity and Kirkhart’s (1995) definition of multicultural validity. Hood coined the term “CRE” in a May 1998 Festschrift, which honored Robert Stake’s influential work in responsive evaluation.
Messick’s seminal work on a unified concept of construct validity underscores the importance of considering the evidence of value implications in score interpretation as a basis for action and the social consequences of score use. This differs from the traditional view of validity as three distinct and interchangeable entities, namely content, criterion, and construct validities. According to Messick (1995), “a unified validity integrates considerations of content, criteria, and consequences into a construct framework for the empirical testing of rational hypotheses about score meaning and theoretically relevant relationships, including those of an applied and a scientific nature” (p. 741). The six aspects of construct validity include content, substantive, structural, generalizability, external, and consequential validity. These aspects should be integrated when interpreting and using assessment data and test scores to make important decisions regarding students and their learning (See Messick 1989, 1995 for a detailed description).

Consequential validity emphasizes the value implications and social impacts of assessment data on learners. For example, if assessment tasks or test items are biased against ELLs and immigrant students, using the resulting data for accountability purposes (such as grade promotion, scholarship competition, placement, and certification) may lead to unintended negative consequences. Such data may not accurately represent the abilities of ELLs and immigrant students. Instead, their lower level of academic achievement may be attributed to language and/or cultural barriers. Assessors’ or teachers’ beliefs and values are often influenced by their own cultural backgrounds, upbringing, past educational experiences, and the broader sociopolitical context (Brown et al., 2009; Pajares, 1992). As Bond (1995) astutely points out, any human assessors or raters can “hold purely prejudicial beliefs that can affect their objective assessments of others” (p. 24).

CRA should surpass the mere demonstration of psychometric properties of data. It should create ample opportunities for meaningful conversations regarding the implementation, communication, interpretation, and utilization of data to inform teachers’ instructional decisions and students’ learning. Such assessment for learning or formative assessment practices are important not only in the mainstream classrooms (Black & Wiliam, 1998; Stiggins, 2002) but also in classrooms where there is a large group of equity-deserving students (Trumbull & Nelson-Barber, 2019). Teachers require support to learn how to utilize assessments to establish an equitable learning culture that empowers every student to thrive despite challenges and obstacles. In today’s complex classrooms, appropriately designed and implemented, assessment can act as a catalyst for advancing equity among all learners.

In an online learning environment, it becomes imperative to examine the complexity of technology as an intervening variable that can either facilitate or hinder the design and implementation of CRA for all learners. This is because historically marginalized students often have lower levels of digital literacy and may lack access to technological devices such as computers, tablets, smartphones, headsets, and microphones, as well as internet connectivity.

To design and implement CRA, educators can consider the CRE framework. Essentially, CRE explicitly highlights the significance of evaluation within a cultural context. It places value on the shared lived experiences and narratives of members within the community. In the CRE framework, cultural differences are perceived as assets rather than deficiencies. The framework includes nine steps, with cultural competence as its core. These nine steps include preparation for
evaluation, stakeholder engagement, identification of the purpose(s) of evaluation, formulation of relevant questions, evaluation of design, selection and adaptation of instruments, data collection, data analysis, and dissemination and implementation of the results (Frierson et al., 2010, as cited in Bryan & Lewis, 2019). Evaluators must possess a range of skills, competencies, and sensitivities, in line with the guiding principles of the evaluation profession.

To promote discussion about fairness and equity, CRE evaluators or CRA assessors must strive to involve stakeholders of various status levels and with differing power and resources. These stakeholders play a crucial role in framing questions and defining the parameters for collecting information about the subject being evaluated (Hood et al., 2015). This aligns with Henning and Lindquist’s (2022) suggestion of employing a team-based approach and inviting experts from the host country to address the cultural competence of evaluators. Such an etic approach to assessment is of paramount importance in today’s linguistically and culturally diverse classrooms. By designing assessments and considering evidence from a minority perspective, the experiences of marginalized groups can be more accurately represented. Teachers should not work in isolation; instead, they should be provided with opportunities to collaborate with colleagues. Such collaboration may include planning, reviewing, and revising lessons together; engaging in co-teaching; providing feedback concerning each other’s practice; co-designing assessment tasks and rubrics, and working with colleagues to moderate student work. This is akin to Japanese Lesson Study (Lewis et al., 2006). In an online environment, a team-based approach to curriculum planning, teaching, and assessment becomes even more important. It creates a supportive and communal space for teachers, particularly for those new to the profession, to develop pedagogical skills and assessment literacy. Additionally, collaboration with colleagues can help teachers feel less isolated, thereby enhancing their social and emotional well-being.

### Culturally Responsive Assessment

In a special issue of the Journal of Negro Education, Johnson (1998) drew on the influential works of Carol Lee (1992) and Gloria Ladson-Billings (1995) in culturally responsive pedagogy to highlight the importance of pedagogical practices that leverage cultural strengths of the learners. Johnson emphasizes that such pedagogy can effectively engage and motivate historically marginalized students, leading to improved academic achievement. Hood (1998) expanded the scope of this argument to include assessment and evaluation, coining the term Culturally Responsive Assessment and Evaluation. Hood asserted that culturally relevant pedagogy necessitates assessment and evaluation strategies that align with its core principles, suggesting that student learning can be “more effectively assessed by using approaches that are also culturally responsive” (1998, p. 189). He also called for a re-evaluation of the cultural values that have traditionally shaped psychometric constructs, challenging prevailing measurement techniques. Hopson (2009) traced the theoretical origins of CRA to Indigenous and minoritized knowledge systems, questioning the dominance of Western, colonizing ones. While assessment experts like Stiggins (1995), Popham (2014), and Shepard (2000) have long advocated for bias-free assessment from the perspective of psychometrics, constructivism, and social constructivism, CRA has gained support in the field of education due to its increased emphasis on equity and social justice. Many countries around the world have engaged in political discourse and implemented deliberate policies to address racial tensions, religious discrimination, violence, radical ideologies, and civil wars. In North America, decolonizing the curriculum has emerged as a key lever in response to the increased recognition of Indigenous and minoritized communities.
CRA intersects with the principles of equity-minded and equity-centered assessment, as proposed by Henning and Lundquist (2022) in the context of higher education. At the core of their framework lies the cultivation of awareness. According to the authors,

> Any attempt to advance educational equity starts with the self. Assessors start by engaging in awareness, learning, and reflection around their identity, privilege, power, and agency, exploring their biases and cultural assumptions. They take personal responsibility for their intention and their impact. Their approach is collaborative, inclusive, relational, and respectful … Cultivating awareness and working to reduce bias and harm are ongoing lenses embedded throughout all approaches in the model.” (p. 187)

Equity-minded assessment practices require assessors and teachers to acknowledge the historical context of oppression and colonization in which assessment is conducted. They must strive to interrupt inequitable systems (local or global), question the decision-makers (policy-makers and administrators), and beneficiaries of assessment (individuals), consider how value is attached to what is assessed (similar to Messick’s consequential aspect of validity), critically analyze how meaning is attached to assessment data and results (also Messick’s consequential aspect of validity), and actively address any systemic barriers that prevent equitable assessment. According to Henning and Lundquist (2022), three approaches enable equity-minded practice: bias-free assessment (a key aspect of assessment literacy in Stiggins, 1995; Popham, 2014; Shepard, 2000), culturally responsive assessment, and socially just assessment. This suggests that both technical and cultural aspects of assessment need to be considered in CRA.

The second component of Henning and Lundquist’s framework is equity-centered assessment, comprising four elements: deconstructed assessment, antiracist assessment, decolonizing assessment, and assessment for social justice. These have become increasingly important given that education systems and assessments of learning outcomes are shaped by a range of cultural, social, political, and technical complexities after a global crisis. The ultimate goal of equity-centered assessment is to place equity at the center of all assessment practices and to utilize assessment data and multiple sources of evidence to dismantle systems of power and oppression. According to Henning and Lundquist (2022), deconstructed assessment intentionally exposes the social structures that perpetuate power imbalances and oppression. Antiracist assessment builds on deconstructed assessment and focuses on identifying and challenging the racist nature of programs, services, policies, and practices rooted in white supremacist assumptions and biases. Decolonizing assessment intentionally incorporates non-Western learning outcomes. Lastly, assessment for social justice needs to intentionally combine the principles of deconstructed assessment, antiracist assessment, and decolonizing assessment, with a primary focus on advancing equity and social justice for those who are historically under-represented and underserved (e.g., Indigenous students and students of color).

Although Henning and Lundquist’s (2022) framework is developed for higher education, I think it is also invaluable in elementary and secondary education. Many of the principles align with authentic assessment or authentic performance assessment and formative assessment, which are student-centered and recognize the value of the teacher’s judgment of students’ work.
assessing students’ work, it is important for teachers to be aware of their own biases, to be equity-minded, and to strive for equity-centered practices.

Designing Culturally Responsive Assessment for Online Learning Environments

Authentic Performance Assessment

During the late 1990s, proponents of educational and assessment reforms encouraged teachers to employ alternative forms of assessment, aiming to capture students’ abilities to apply their knowledge and skills in real-world contexts (Newmann et al., 1996; Pellegrino et al., 2001; Shepard et al., 2005). Performance assessments, also known as performance-based tasks, emerged as a pedagogically sensitive and valid method for measuring students’ learning progress and outcomes. This approach was believed to offer more accurate data or evidence on students’ higher-order thinking skills and other essential 21st century competencies. During that period, portfolios emerged as a commonly used form of performance assessment, wherein students could select and showcase their best work and to reflect on their learning. However, when portfolios were adopted as a large-scale high-stakes assessment in several U.S. states (e.g., Vermont, Kentucky, Maryland, and California), numerous controversial issues arose due to variations in teachers’ implementation practices. Concerns arose regarding the following: teachers’ treating of the tasks as if they were summative tests or exams and hence teaching to them; the authenticity of the student’s work produced, when excessive assistance was offered by teachers and/or parents; variations in revision opportunities across different classes and schools and, the allotted time for students to complete the tasks. Numerous measurement experts expressed concern regarding the reliability and validity of using students’ portfolios to make high-stakes decisions. For instance, Holland (2007) states that portfolios “may be a useful motivational device for a classroom teacher working with individual students, but judging the worth of collected reports about disparate topics is inherently too subjective to be a basis for large-scale assessment.” This indicates that portfolios are more suitable to be used as formative assessment.

The term “authentic” was initially introduced in the context of assessment by Archbald & Newmann (1988). Newmann et al. (1996) further unpacked the construct of authentic achievement to include three broad criteria: construction of knowledge, disciplined inquiry, and value beyond school. They argued that “construction of knowledge through disciplined inquiry to produce discourse, products, or performance that have value beyond success in school, can serve as a standard of intellectual quality for assessing the authenticity of student performance” (p. 287). Grant Wiggins (1989) is credited with coining the term “authentic assessment” and providing a definition for it. According to Wiggins, authentic assessments seek to “replicate the challenges and standards of performance typically faced by writers, business people, scientists, community leaders, designers, or historians” (p. 704). The authenticity of performance assessments or performance-based tasks lies in their open-endedness, genuine nature, and complexity, as well as the real-world problems embedded within them, which afford students opportunities to apply their knowledge and skills to solve problems in a real-world context (Koh, 2017; Wiggins, 1989). In essence, authentic tasks are set up at a high-level cognitive demand. Thus, they engage students in high-level thinking. Adapting from the works of Newmann et al. (1996), Anderson & Krathwohl (2001), Koh & Luke (2009) and Koh (2011a & 2011b) have proposed a framework known as the Authentic Intellectual Quality (AIQ) framework. This framework consists of two sets of criteria that serve as guidelines for teachers to evaluate the
quality of classroom assessment tasks, design authentic assessments, and judge the quality of students’ work in relation to these tasks.

Newmann et al.’s (1996) influential research on the implementation of authentic assessment in Chicago public schools during high school, yielded significant findings. Their study demonstrated that when teachers in disadvantaged schools assign authentic performance tasks that require higher-order thinking, in-depth understanding, elaborated communication, and making meaningful connections to students’ real lives, students produce more intellectually complex work. This finding suggests the notion that what you test is what you get. However, Lee (1998) argued that Newmann et al.’s research did not establish a direct correlation between authentic assessment, pedagogy, and improved academic outcomes for underachieving minority students. She emphasized the need to address equity issues, which encompass more complexities beyond the intellectual rigor and contextualization of assessment tasks. Both Lee (1998) and Hood (1998) advocated for the development of culturally responsive performance assessments as a means of fostering equity for students of colour.

Given the growing diversity of learners in PreK-12 and higher education settings, equity has emerged as one of the four fundamental principles—equity, diversity, inclusion, and accessibility guiding policy initiatives, research priorities, and educational practices worldwide. In certain countries like Canada, the United States, Australia, and New Zealand, decolonization has been recognized as an intentional approach to honor and uphold the inherent rights of Indigenous peoples. Over the past decade, substantial attention has been devoted to academic discourse and policymaking centered on equity in education. For instance, the Assessment Institute in the United States has advocated for evidence-based assessment practices that foster and bridge learning gaps in higher education (Hansen & Renguette, 2021). However, there remains a limited understanding of how to effectively design and implement equity-centered and socially just assessments, such as CRA, in contemporary classrooms, particularly in online learning environments. Therefore, it is crucial to revisit the historical foundations of CRA and reflect on its role in ensuring reliable, valid, and equitable assessment in online environments for equity-deserving students.

Key Considerations for Culturally Responsive Online Assessment

Outlined below are seven key aspects to consider when designing and implementing culturally responsive online assessments for equity-deserving students.

**Narrative and Contextual – Recognizing Indigenous Ways of Knowing, Doing, and Being**

Contextualization plays a vital role in authentic performance assessment (Cumming & Maxwell, 1999). Authentic assessment entails the use of rich, open-ended performance-based tasks that are situated in real-world scenarios. Such tasks, including projects, designs, discourses, and performances, are often perceived by students for having “utilitarian, aesthetic, or personal value” (Newmann et al., 1996, p. 284). To be effective, the problems or challenges embedded in authentic tasks should be culturally specific, enabling students to establish connections between their learning within a subject or across multiple subjects (transdisciplinary learning) and the real world. It is widely acknowledged that authentic tasks foster greater intellectual engagement compared to conventional standardized tests as they elicit higher-order thinking skills and 21st
century competencies (Koh & Luke, 2009). As a result, authentic tasks can be a powerful tool for advancing equity in today’s increasingly diverse and complex classrooms.

In authentic performance assessment, students should have agency in determining how they present their work. Providing multimodality options, such as narratives, digital stories, metaphors, texts, graphics, charts, graphs, audio, images, animations, and video, ensure that all students, regardless of their learning styles and available resources, have an equitable opportunity to showcase their work. Some researchers argue that culturally responsive performance assessments enable teachers to nurture students’ cultural identity and enhance their authentic learning experiences by providing opportunities to engage in intellectual work and discourse. In an online learning environment, teachers can encourage students to reflect on their learning through avenues like blogs, learning logs, journals, or e-portfolios. ELLs and Indigenous students, in particular, may choose to share their personal stories, as well as key aspects of their respective cultures.

Recognizing the significance of Indigenous ways of knowing, doing and being is more crucial than ever before. In the context of mathematics, as an example, the assessment of students’ performance can perpetuate social and cultural injustices when the comprehension and practice of mathematics are typically defined through a Western perspective of mathematical knowledge. This implies that school mathematics should consider the learner’s cultural knowledge, including Indigenous ways of understanding and performing mathematics. Nasir et al. (2008) have long contended that mathematical knowledge is “an inherently cultural activity” (p. 191), and that failing to address inequalities arising from race, identity, and power in the teaching and learning of mathematics may benefit certain groups of learners and disadvantage others. It underscores the importance of teachers’ abilities to understand the sociocultural aspects of mathematics learning and consider the “social processes underlying the teaching and learning of mathematics” (Nasir & de Royston, 2013) in linguistically and culturally diverse classrooms. Many assessment experts (e.g., Moss et al., 2006) have long called for a consideration of the influences of learning environments and the larger sociocultural contexts.

**Language and Translation Issues of Assessment**

In addition to honoring cultural differences, CRA practices should also consider the use of inclusive language. In the late 1990s and early 2000s, a substantial body of empirical research emerged, focusing on translation effects in national and international assessments involving English- and French-speaking students’ achievement in mathematics and science. These assessments included the Canadian Student Achievement Indicators Program, the Trends in International Mathematics and Science Study, and the Program in International Student Assessment (e.g., Ercikan, 1998; Ercikan et al., 2004; Ercikan & Koh, 2005). Cutting-edge measurement methods, such as multiple-group confirmatory factor analysis, Item Response Theory (IRT), and Non-IRT differential item functioning (DIF), were employed during that time to examine the comparability of different language versions of tests and items (e.g., English versus French). DIF refers to items that perform differently for different groups of students (e.g., based on gender, ethnicity, language, and culture) with the same ability level. Establishing measurement invariance at both the test and item levels is considered crucial evidence for construct validity in most cross-cultural research (Koh & Zumbo, 2008; Zumbo & Koh, 2005). Researchers discovered DIF items even when comparing international achievement data between
Designing Culturally Responsive Online Assessments

American English-speaking and British English-speaking students. This measurement invariance approach holds value when translating assessment and survey items to another language for use with immigrant or ethnic minority students. It can also be employed to examine measurement invariance between different groups of learners, such as male versus female students, or Indigenous versus non-Indigenous students (e.g., Salazar-Fernández et al., 2020).

Cultural Validity

Language plays a pivotal role in the field of assessment, as language differences among students can impact how they interpret test items or assessment tasks. Trumbull & Solano-Flores (2011) emphasize the significance of language in considering cultural validity in assessment. For instance, it is important to view Indigenous students’ heritage languages as assets when assessing their learning. Assessments designed for Indigenous students should also draw on the emotional and spiritual dimensions of their learning. Trumbull & Nelson-Barber (2019) advocate for teachers to employ informal formative assessments, such as reflection, journal writing, and oral questions, within Indigenous classrooms. They also cautioned against the translations of English test questions or assessment tasks into students’ native languages, as this can introduce potential challenges. Advanced measurement approaches may not adequately address inherent biases stemming from language and cultural differences, making it highly unlikely to achieve full comparability between two language versions of tests or assessments. Therefore, it would be more effective to develop assessment tasks in students’ native language. As Wiliam (2010) aptly points out, defining the construct of interest in a manner appropriate for ELLs, Indigenous students, students of color, or students with special needs is a more viable approach to addressing equity than modifying the assessment itself. Messick (1994) advocates for a construct-centered approach to performance assessment (authentic assessment), suggesting that it “would begin by asking what complex of knowledge, skills, or other attributes should be assessed presumably because they are tied to explicit or implicit objectives of instruction or are otherwise valued by society” (p. 16). Ultimately, the careful design or selection of assessment tasks and the development of scoring rubrics should be driven by the learning processes and outcomes valued by the community. Messick (1994) further notes: “Focusing on constructs also alerts one to the possibility of construct-irrelevant variance that might distort the task performance, its scoring, or both” (p. 16).

Moderation Practice

While implementing culturally responsive online assessments, it is important to ensure that teachers’ judgements of student work are defensible. Therefore, proper online moderation practice needs to be put in place. From a sociocultural lens, moderation is a social practice. It involves teachers working together to discuss and reach consensus on their judgements of the quality of student work using agreed-upon performance standards and success criteria (Adie & Klenowski, 2016). Such an approach enables teachers to reflect on and address personal biases in their assessments as human judgements are often mediated by one’s own beliefs, values, attitudes, and cultural context. Scholars (e.g., Koh, 2014; Smaill, 2020) have recognized that teachers’ involvement in social moderation practice serves as a powerful mechanism for professional growth and development, especially in building assessment capacity.
Analysis of Response Process Data

Over the past two decades, new measurement approaches have emerged, including think-aloud protocols, cognitive diagnostic models, and analysis of response process data, which offer multiple sources of evidence for construct validation. While these approaches are primarily used in large-scale assessments and require advanced statistical skills for meaningful data analysis, they provide a nuanced understanding of students’ cognitive processes, emotions, motivations, and behaviors when taking tests or completing performance tasks (Ercikan et al., 2020; Hubley & Zumbo, 2011). Drawing on Messick’s (1989) progressive matrix model of unified validity theory, Hubley & Zumbo (2011) highlight that “if test developers [assessment designers] and users want measures to have personal and social consequences and impact, then it is critical to consider the consequences and side effects of measurement in the validation process itself” (p. 219). Therefore, it is crucial to collect and analyze response process data. Ercikan & Pellegrino (2017) define response processes as the “thought processes, strategies, approaches, and behaviors of examinees when they read, interpret, and formulate solutions to assessment tasks” (p. 2). This data can be captured through various methods, such as verbalizations, eye tracking, physiological measures, response times, and computer clicks. Technological advancements have facilitated the collection of response process data, offering a detailed analysis of test takers’ interactions with test items or assessment tasks, as well as their experiences, emotions, behaviors, and engagement, particularly in digital assessments (Maddox, 2023). Analyzing students’ response process data serves as an additional source of evidence for construct validity and can also inform the design of CRA. With the increasing use of technology, it is essential to develop teacher-friendly approaches that enable the timely sharing of students’ process data and learning analytics with teachers. This will help teachers gain a better understanding of how their students learn, and the data can inform instructional plans and the selection of appropriate pedagogical and assessment approaches to meet the needs of diverse learners. Such assessments are formative and play a pivotal role in supporting student learning, known as—assessment for learning or formative assessment (Black & Wiliam, 1998). According to Ercikan et al. (2020), “Central to this role [formative assessment] is information about how students engage with assessment tasks, what processes they follow, what kinds of errors [misconceptions] they make in coming up with their solutions, all focusing on response processes rather than final solutions” (p. 2).

Technological/Digital Literacy

In the digital and AI era, the progress of technology has made it essential for PreK-12 teachers and college/university instructors to familiarize themselves with new technologies and AI tools. Even though most students have returned to in-person learning, there is still room for designing culturally responsive online assessments that complement students’ learning and foster their information and communication technology (ICT) skills and digital literacy. As Warschauer et al. (2010) aptly point out, “Access [to technology] alone will not overcome inequity in use and outcomes” (p. 215). The key lies in utilizing Internet-connected computers to bring “the wider world into the classroom” (p. 215). For example, students can access online reading materials and resources, as well as participate in virtual field trips, museum visits, and STEM activities to explore different cultures, communities, and natural wonders across the globe. Such contextualization offers every student authentic experiences, thereby enhancing their interest and motivation to learn (Koh, Chapman & Lam, 2022).
Building Teacher Capacity in Culturally Responsive Online Assessment

Teacher professional development and initial teacher preparation need to integrate cultural approaches to learning and assessment into workshops and coursework. Trumbull & Nelson-Barber (2019) observe that “classroom assessment can also fail Native [Indigenous] students, when teachers do not have a sophisticated understanding of these students’ culture-based approaches to learning or when they feel pressured by state policies to standardize instruction and assessment,” (p. 2). Teachers who lack ICT literacy often experience high levels of technostress, which can have a negative impact on their assessment practices in online environments. Research has shown that teachers’ technostress can hinder their adoption of new technologies in practice (e.g., Joo et al., 2016). Therefore, professional development programs on the design and implementation of culturally responsive online assessments should focus on enhancing pre-service and in-service teachers’ capacity to examine the intersections between assessment, culture, and technology, with equity as the center of all conversations and practices (Henning & Lundquist, 2022). In the digital age, teachers need to stay up to date concerning new technologies including virtual reality, augmented reality, machine learning, learning analytics, ChatGPT, and other AI tools. Moreover, teachers should receive adequate support to effectively utilize these new technologies in online environments, promoting culturally responsive assessments that are fair and equitable for all students.

Conclusion

In conclusion, this article provides a brief review of CRA. The author highlights key considerations in the design and implementation of culturally responsive online assessments for equity-deserving students. While the review may not be exhaustive, it aims to shed light on equity and assessment during global crises such as the COVID-19 pandemic. In their editorial comment on what counts as evidence and equity in the American Educational Research Association’s flagship journal Review of Research in Education, Luke et al. (2010) caution that, “To address questions of equity require rich, interpretive, and evolving sciences, not a narrow technical approach that invites capture by particular doctrinal and generic approaches to systems reform, public policy, and institutional governance” (p. xv). This suggests that further empirical research is necessary to provide meaningful evidence on the role of culturally responsive online assessment in advancing equity for all students.
References


https://doi.org/10.1080/10627197.2020.1804353

https://doi.org/10.1207/s15327574ijt0501_3


https://doi.org/10.1002/au.30258


https://doi.org/10.2307/2668188


https://doi.org/10.1007/s11205-011-9843-4

https://doi.org/10.2307/2668186

https://doi.org/10.1016/j.compedu.2015.12.004

https://doi.org/10.1177/109821409501600101


Statistics Canada. (2022). *Immigrants make up the largest share of the population in over 150 years and continue to shape who we are as Canadians.* https://www150.statcan.gc.ca/n1/daily-quotidien/221026/dq221026a-eng.htm


