

IMAGINE Research of Transformational Assessment

Anthony J. Onwuegbuzie¹

¹University of Johannesburg

Elena Forzani²

²Boston University

Sandra Schamroth Abrams³

³University of South Africa, University of Sheffield

Abstract

In the context of studying human beings, at the macro level, assessment is the process of documenting, analyzing, and interpreting empirical data to provide evidence regarding an observable entity concerning humans that emerges cognitively, affectively, physically, metaphysically, and/or spiritually, and which is internal or external to the person or group being observed. In this article, we argue that assessment does not belong exclusively to either the qualitative or quantitative research tradition; rather, it is central to both traditions, as well as to mixed methods research. Moreover, in every research study—whether representing the qualitative, quantitative, or mixed methods research tradition—some form of assessment is used. Further, assessment not only represents a social constructionist and/or a social constructivist act, but also, and even more importantly, it represents a political act. This article features a discussion about the systemic, historically disempowering nature of assessment that privileges an artificial and racist norm, and silences and punishes participants who become othered by traditional research and assessment paradigms. Thus, we invite researchers—regardless of their methodological experiences and orientations—to embrace an **Integrative Mixed Methods Antiracist Groundwork for Investigating and Nurturing Equity** (IMAGINE; S. S. Abrams et al., 2021, 2022; Onwuegbuzie et al., in press-b), a research and evaluation meta-framework that we use here to transform assessment to support more equitable and empowering research. Utilizing critical dialectical pluralism (Onwuegbuzie et al., in press-a; Onwuegbuzie & Frels, 2013), which underscores the inclusion of participants and their voices at every stage of the research process—from conceptualization, to dissemination, to utilization—we focus on research that honors participants as co-researchers, co-ideators, and co-decision makers, who, through their involvement in the research, become activists striving to identify problems and to effect change that they envision in local and/or global contexts.

Keywords: mixed methods research, critical dialectical pluralism, antiracist research framework, assessment in education, culturally responsive assessment, critical assessment practices approach, equity-oriented assessment

Introduction

According to the Collins English Dictionary (2021), *assessment* (circa 1530–1540; Middle English) firstly is defined as “a consideration of someone or something and a judgment about them” (1). Similarly, Merriam-Webster (n.d.) firstly defines assessment as “the action or an instance of making a judgment about something: the act of assessing something: APPRAISAL” [emphasis in original] (¶1). However, the

Cambridge Dictionary (2021a) likely provides the most comprehensive definition, by defining assessment as “the act of judging or deciding the amount, value, quality, or importance of something, or the judgment or decision that is made” (¶1). Therefore, assessment is a term that transcends research, in general, and research approaches and paradigms, in particular. Moreover, assessment does not belong

to either the qualitative research tradition or the quantitative research tradition; rather, it is a central element of both traditions, as well as of mixed methods research. In fact, we argue that in every single research study—whether representing the qualitative, quantitative, or mixed methods research tradition—some form of assessment takes place.

Furthermore, given that assessment includes judgment and value-based decisions, assessment can be—and historically has been—a hotbed for discriminatory and hierarchical practices that perpetuate “scientific racism” (Saini, 2019, p. 29). Looking to dismantle and to flatten such hierarchies and to reframe and to reform research and evaluation, we propose a meta-framework to support more equitable and empowering research. After situating assessment historically, we introduce the IMAGINE movement—an Integrative Mixed methods Antiracist Groundwork for Investigating and Nurturing Equity—its methodological framing and its practical applications vis-à-vis a Critical Assessment Practices approach (CAPS; Forzani, Dobbs, et al., 2024).

Brief History of Western-Based Educational Assessment

Unfortunately, assessment in both quantitative research and qualitative research—and thus, mixed methods research—has a sordid past. With respect to qualitative research, for example, the first half of the 20th century was characterized by colonial forms of inquiry (Denzin & Lincoln, 2011; Erickson, 2018; Patel, 2016; Vidich & Lyman, 1994, 2000). These research studies took place in foreign settings, where White Western researchers (e.g., Lone Ethnographers; Rosaldo, 1989) conducted fieldwork that involved *so-called* “objective,” “imperialist,” “monumentalist,” and “timeless” assessments of the culture, customs, habits, and religions of the “Other”—which predominantly involved those from what now is referred to as the Global South (i.e., lower-income countries) (Denzin & Lincoln, 2011; Vidich & Lyman, 2000). Findings and, even more importantly, interpretations, stemming from many of these

research studies, were extremely harmful (see, for e.g., Malinowski’s [1967] observations of his field experiences in New Guinea and the Trobriand Islands from 1914-1915 and 1917-1918) because they were driven by what can be called *White racial framing*. This framing represents a formal and informal racial hierarchal structure comprising White people at the top and Black (e.g., African American) people at the bottom (e.g., adopting a Eurocentric framework). Such a framing generally consists of socially constructed characteristics (e.g., negative racial stereotypes) that are designed to create and to sustain both White superiority and the inferiority of other groups (Feagin & Cobas, 2008; Taylor, 2006).

With regard to quantitative research, the turn of the 20th century witnessed the onset of the era of standardized testing. This era began with the development of the first standardized admissions test in 1901 by a group of U.S. colleges—namely, the College Entrance Examination Board—that assessed how well students were prepared for college-level coursework. Soon thereafter, in 1905, Alfred Binet introduced the first modern standardized test of intelligence, which directly assessed students in order to identify who needed educational assistance (Brink, 2011). Two years later in 1907, Karl Pearson first used a 7-point scale in research on intelligence (circa March 27, 1857–April 27, 1936) (McReynolds & Ludwig, 1987). Pearson, an English mathematician and biostatistician who is credited with establishing the discipline of mathematical statistics, founded the first university statistics department at University College, London in 1911. Unfortunately, Pearson also was a vocal proponent of *social Darwinism* (i.e., applying biological concepts of natural selection and survival of the fittest to social science fields [e.g., sociology]; Williams, 2000), *eugenics* (i.e., beliefs and practices that exclude people and groups who are judged to be inferior and promote those judged to be superior; Galton, 1904), and *scientific racism* (i.e., the pseudoscientific belief that empirical evidence exists to justify racism, racial superiority, and racial inferiority; Saini, 2019). These and other quantitative assessments developed in this era reflected a top-down (i.e., hierarchical),

elitist, highly context-specific, culture-specific, normative-based, product-based, independent, competitive, punitive, inauthentic, and inequitable approach to assessment development—wherein the most important partners—the participants (or, in this case, the subjects)—were omitted from the development process. Consequently, these measures were ethnocentric and gender centric, yielding extremely flawed and systemically racist assessments that were, and continue to be, used across the social and behavioral sciences—including education—and the health sciences. For example, flawed assessments have led repeatedly to the conclusion that there are racial differences in intelligence that are substantially genetic in origin (Daley & Onwuegbuzie, 2011, 2020; Onwuegbuzie & Daley, 2001). As concluded by Daley and Onwuegbuzie (2020),

in this post-truth era, with respect to intelligence research, what we have is a strong relationship between two weak phenomena (i.e., race and intelligence), one of which — intelligence — is reported to be measurable with IQ tests that happen to correlate with socioeconomic status and that represent a narrowly defined set of cognitive skills which, not surprisingly, predict similarly defined academic skills and, therefore, occupational success and wealth, which, in turn, predict intelligence as represented by an IQ score. Flawed constructs, flawed instruments, and flawed relationships yield flawed inferences and flawed educational and social policies. (p. 395)

Figure 1 presents an overview of the history of quantitative, qualitative, and mixed methods research from the 20th century onwards, alongside a history of educational assessments associated with these three research traditions for the same time period. This figure was adapted from “Onwuegbuzie, Forzani, and Abrams (2022). The last column of this figure reveals that, in general, qualitative assessments, quantitative assessments, and combinations thereof have not kept pace with the evolution of these three traditions, with quantitative assessments remaining the dominant form of assessment.

Moreover, throughout its history, assessment development practices have led to assessments with questionable reliability, internal validity,

external validity, internal credibility, external credibility, objectivity, trustworthiness, dependability, confirmability, transferability, and/or authenticity, which, in turn, have exacerbated obfuscation in meaning making, thereby promoting maleficence instead of beneficence. These practices have led to dire educational outcomes for a significant proportion of students representing primary, secondary, and tertiary education—especially for those from Global South countries and territories (including regions within the Caribbean, Latin America, Africa, Oceania, Pacific Islands, the Middle East, Eastern Europe and the developing countries in Asia)—as well as minority and indigenous populations (i.e., people of color) from Global North countries and territories (i.e., Europe, North America, Australia, Canada, Russia, Israel, Japan, New Zealand, Singapore, South Korea, and Taiwan). Unfortunately, these dire educational outcomes prevail today (see, for e.g., Daley & Onwuegbuzie, 2020) as the result of a lack of several major areas of validity/legitimation evidence—most notably, the following three major areas of validity/legitimation evidence that form part of Onwuegbuzie et al.’s (2009) meta-validation model, which, in turn, stem from Messick’s (1989, 1995) conceptualization of validity: *substantive validity* (i.e., “Assesses evidence regarding the theoretical and empirical analysis of the knowledge, skills, and processes hypothesized to underlie respondents’ [responses]” [p. 202]), *generalizability* (i.e., “the extent that meaning and use associated with a set of ...[responses] can be generalized [or transferred] to other populations” [p. 202]), and *outcome validity* (i.e., “the meaning of...[responses] and the intended and unintended consequences of using the instrument” [p. 202])—or what some authors refer to as *consequential validity* (e.g., Brewer et al., 2014; Hitchcock et al., 2015). Therefore, the time is ripe—nay, overdue—for the development of assessment systems that simultaneously maximize these areas of validity/legitimation evidence. We contend that such assessment would be beneficial to those being assessed in both the Global South (e.g., the Caribbean) and the Global North.

Figure 1

History of Quantitative, Qualitative, Mixed Methods Research, and Western-Based Educational Assessment: Twentieth Century to Present Day

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
1900–1929	<p><i>Formal emergence of the social and behavioral sciences:</i></p> <p>Classical positivism: introduced by Auguste Comte (French Philosopher)</p> <p>Logical positivism (circa 1920s): originated in the Vienna Circle, a group of European Scholars</p> <p>Birth of hypothetico-deductive model</p> <p>Development of basic statistical and anthropological methods</p>	<p><i>Moment 1: Traditional:</i></p> <p>Many researchers who rejected (logical) positivism embraced the qualitative research paradigm.</p>	<p><i>Formal emergence of the social and behavioral sciences</i></p>	<p><i>Qualitative: Era of Colonial forms of qualitative research:</i></p> <p>Qualitative research studies in foreign settings wherein White researchers (e.g., <i>Lone Ethnographers</i>) conducted fieldwork that involved so-called “objective,” “imperialist,” “monumentalist,” and “timeless,” assessments of the culture, customs, habits, and religions of the “Other”.</p> <p>The Chicago school established the assessment of human group life.</p> <p><i>Quantitative: Era of standardized testing:</i></p> <p>A team of U.S. colleges—namely, the College Entrance Examination Board—developed the first standardized admissions test to assess whether students were prepared for college-level course work (1901).</p> <p>First course in educational assessment that was taught by Thorndike at Columbia in 1902 (Meyer, 1965)</p> <p>Alfred Binet introduced the first modern standardized test of intelligence, which directly assessed students in order to identify students who needed educational assistance (1905).</p> <p>Karl Pearson is deemed to be the first investigator to employ ratings—namely, a 7-point scale—in research on intelligence (1907).</p> <p>First definition of true score in 1910 (Brown, 1910).</p> <p>National Council of Education published a major report on standards and tests for assessing school efficiency (1913) (cf. Strayer, 1913).</p> <p>By World War I, standardized testing was standard practice: aptitude quizzes called Army Mental Tests were conducted to assign U.S. servicemen jobs during the war effort.</p> <p>World War I yielded a surge in psychological testing as thousands of U.S. recruits are screened for intellectual and emotional functioning (1914).</p> <p>The multiple-choice test was invented to combat the rise in student population in the United States (1915).</p> <p>Stanford-Binet IQ test was created (1916).</p> <p>Swiss psychiatrist Hermann Rorschach published his famous monograph, <i>Psychodiagnostics</i>, which led to the development of the Rorschach Inkblot Test to examine a person’s personality characteristics and emotional functioning (1921). This test subsequently was used to assess students in school settings (e.g., mental tests, childhood and adolescence, educational psychology vocational guidance; cf. Hertz, 1934).</p> <p>Rugg published a four-part paper on rating scale methodology (1921-1922).</p> <p><i>Journal of Educational Measurement</i> devoted several issues in 1921 to a symposium on scientific assessment of intelligence.</p> <p>The Scholastic Aptitude Test (SAT) was developed in 1926 by the College Board, which assessed knowledge of vocabulary and basic mathematics.</p> <p>Carl Spearman published a two-factor theory of intelligence in which he postulated the existence of a general intellectual ability factor and specific components of that general ability (1927).</p> <p>Louis Leon Thurstone, a U.S. pioneer in the fields of psychometrics and psychophysics, developed the Thurstone Scale to assess attitudes towards religion (1928), which was subsequently used in educational constructs.</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
1930-1949	<p><i>Traditional period:</i></p> <p>Discrediting of logical positivism</p> <p>Early forms of postpositivism</p> <p>Further extensions to the hypothetico-deductive model</p>	<p><i>Moment 1: Traditional</i></p>	<p><i>Traditional period:</i></p> <p>Uncontroversial but limited use of mixed methods research</p>	<p><i>Qualitative: Era of Chicago school:</i></p> <p>The Chicago school continued promoting the assessment of human group life.</p> <p><i>Quantitative: Era of standardized testing and study of eye movement:</i></p> <p>Miles Tinker and his collaborators began using photographic techniques to study eye movement of readers (1930).</p> <p>G. T. Buswell created the first non-contact device registering eye movements to assess the reading and watching of images (1935).</p> <p>Rensis Likert, psychologist, developed the Likert-format scale to assess people's attitudes, opinions, and perceptions (1932).</p> <p>The first automatic test scanner was developed (January 1, 1936).</p> <p>Kuder and Richardson (1937) published a seminal article on test reliability.</p> <p>By 1938, more than 4,000 psychological tests were in print.</p> <p>SAT is normalized to make test scores as fair as possible (1941).</p> <p>Minnesota Multiphasic Personality Inventory was published to assess personality (1943).</p> <p>Government-funded schools began using written examinations that were introduced first in Boston (1945).</p> <p>Cronbach (1947) introduced several different kinds of reliability coefficients.</p> <p>The first version of the Wechsler Intelligence Tests for children was published (1949).</p>
1950-1959	<p><i>Postpositivist era:</i></p>	<p><i>Moment 2: Modernist or golden age</i></p> <p>Attempt to make qualitative research (e.g., grounded theory) as rigorous as quantitative research; causal narratives were central; many texts attempted to formalize qualitative research; new interpretive theories emerged (e.g., ethnomethodology, critical theory, feminism, phenomenology).</p>	<p><i>Postpositivist era</i></p> <p>Campbell and Fiske (1959) formalized the practice of using multiple research methods by introducing the concept of triangulation.</p>	<p><i>Qualitative: Era of the modernist ethnographer and sociological participant observers:</i></p> <p>"Rigorous" qualitative assessments were taken of social processes. Postpositivism was the paradigm that greatly influenced qualitative assessment practices via new interpretive frameworks (e.g., phenomenological, ethnomethodology, critical theory, feminism). In particular, these qualitative researchers applied Campbell and Stanley's (1963) conception of internal validity and external validity to constructionist and interactionist notions of assessment in an attempt to make qualitative research as rigorous as quantitative research. These qualitative research studies often were characterized by assessments that were based on a combination of open-ended and quasi-structured interviewing schedules, as well as by participant observations.</p> <p><i>Quantitative: Era of standardized testing and attribute assessment:</i></p> <p>Louis Guttman developed the Guttman scale, which is a single (unidimensional) ordinal scale for the assessment of the attribute (1950).</p> <p>Publication of the first standards for educational and psychological assessment (American Psychological Association [APA], 1954).</p> <p>The first version of the Wechsler Adult Intelligence Tests was published (1955).</p> <p>The term Artificial Intelligence (AI) was coined by John McCarthy, an American computer scientist and cognitive scientist (1956).</p> <p>American College Testing (ACT) was developed as a competitor to the SAT (1959).</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
1960-1969	Publication of Thomas Kuhn's (1962) book	<p><i>Moment 2: Modernist or golden age</i></p> <p>Publication of Thomas Kuhn's (1962) book.</p>	<p><i>Postpositivist era:</i></p> <p>Emergence of multimethod designs</p> <p>Webb et al. (1966) conceptualized the use of multiple methods (i.e., multiple operationalism).</p> <p>Promotion of the use of quantitative and qualitative methods in social research.</p>	<p><i>Qualitative: Era of the cultural romantics:</i></p> <p>Qualitative researchers, as cultural romantics, conducted qualitative assessments via an ironic and tragic view of society (Denzin & Lincoln, 2011).</p> <p>Farrah et al. (1968) developed the Self-Concept and Motivation Inventory (SCAMIN). Although it has been subsequently used as a quantitative assessment (Davis et al., 1988), it was originally developed as a qualitative assessment, serving as a graphic method for use with children (Farrah et al., 1968). This graphic method involves children being shown different drawings of faces, comprising one face with a happy smile, a second face with a straight line for a mouth, and a final face with a mouth turned downward to depict sadness. Children are asked to select which face best characterizes their feelings about certain experiences, such as how they feel about school. For example, children can be asked to report which face belongs to each of their classmates, which would yield a SCAMIN drawing.</p> <p><i>Quantitative: Era of personality assessment</i></p> <p>Crespi developed the Stapel Scale (1961).</p> <p>Meyers Briggs Type Indicator (MBTI) was published (1962).</p> <p>Warren T. Norman published his first article on the Big Five Personality Factors (1962).</p> <p>Criterion-referenced and norm-referenced tests were coined by Robert Glaser, a U.S. educational psychologist (1963).</p> <p>The American Psychological Association (APA), American Educational Research Association (AERA), and National Council on Measurement in Education (NCME) jointly revised the standards for educational and psychological assessment (APA, AERA, & NCME, 1966).</p> <p>Fee waivers to take the SAT started to be offered to all eligible students who could not afford the registration fee (1969).</p> <p>Automated facial recognition was pioneered (1960s).</p>
1970-1979	Emergence of the causal model of explanation	<p><i>Moment 3: Blurred genres</i></p> <p>Qualitative researchers had a full arsenal of paradigms, methods, and strategies; computers came to the fore to aid qualitative analyses; new approaches surfaced (e.g., poststructuralism, neopositivism); several qualitative journals emerged; naturalistic, postpositivist, and constructionist paradigms gained power.</p>	<p><i>Diversification of and advances in methodologies in the human sciences:</i></p> <p>Denzin (1978) outlined how to triangulate methods.</p>	<p><i>Qualitative: Era of blurred genres in qualitative assessment:</i></p> <p>Qualitative assessment reflected blurring of boundaries between the social sciences and the humanities, with semiotics and hermeneutics playing an important role.</p> <p>Ecological maps or, more simply, ecomaps (i.e., eco-maps; also known as ecograms), developed in 1975 by Hartman (cf. Hartman, 1978, 1995), are graphical representations that illustrate all of the systems involved in an individual's life. Although eco-maps were developed to be used in individual and family counseling within the social work and nursing professions, they can be used to study the field of education, for example, to record information of learning experiences and to show how these interactions support or hinder a student (Bennett & Grant, 2016). Information about important interactions in a student's life can help teachers and administrators to understand students in ways that might not be revealed via educational interactions (Bennett & Grant, 2016). This first-hand knowledge of a student's strengths or weaknesses can help teachers and administrators learn how to individualize student development, and can be an important part of students' performance, as well as their personal and professional growth (Bennett & Grant, 2016).</p> <p>Formalization of the ethnographic interview by Spradley (1979), wherein the assessor is "more collaborative and informal" with the assessed and "does not try to maintain an 'objective' or formal distance" from the assessed (Franklin & Jordan, 1995, p. 283).</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
1980-1989	Paradigm wars	<p><i>Moment 3: Blurred genres</i></p> <p><i>Moment 4: Crisis of representation:</i></p> <p>Research and writing became more reflexive and led to questions about issues of gender, race, and class; new models of truth, representation, and method were sought; issues such as validity, reliability, and objectivity re-emerged as being problematic; triple crises of representation (i.e. qualitative researchers can no longer directly capture lived experience), legitimation (i.e., makes problematic the traditional criteria for evaluating and interpreting qualitative research), and praxis (i.e., involves asking whether it is possible to effect change in the world if society is only and always a text).</p> <p>Emergence of newer paradigms such as constructivism that led to <i>paradigm wars</i>.</p>	<p>Paradigm wars</p> <p>Continued development of rationale for the use of mixed research.</p>	<p><i>Quantitative: Era of criticism of standardized assessment</i></p> <p>APA, AERA, and NCME jointly published the second revision of the standards for educational and psychological assessment (APA, AERA, & NCME, 1974). This edition included standards for employment and college admissions testing and addressed test development, test use, and reporting, expanding the focus beyond only test development and documentation.</p> <p>The Joint Committee on Standards for Educational Evaluation, an American-/Canadian-based Standards Developer Organization, was formed in 1975, to develop evaluation standards and to improve the quality of standardized assessment.</p> <p>An increased skepticism prevailed about the efficacy and usefulness of student achievement and intelligence tests for schools, as well as perceived testing monopolies. These criticisms yielded some important legislative changes concerning the testing industry—particularly, the Truth-in-Testing law that was passed by New York in 1979, which required sponsors and manufacturers of college admissions examinations to offer test takers the right, at the time scores are reported, to obtain copies of the test along with their answer sheet and a key to the correct responses.</p> <p>American psychologist, Charles Egerton Osgood, developed the semantic differential scale to assess the <i>connotative meaning</i> of emotional attitude towards various matters (1979).</p> <p><i>Qualitative: Era of crisis of assessment:</i></p> <p>Qualitative assessment became more reflexive. Conflicts emerged between assessment and reporting.</p> <p>Formalization of the participant observation by Spradley (1980), which “vary along a continuum that encompasses two dimensions—observation and participation” (Franklin & Jordan, 1995, p. 289), wherein assessments can be obtained via many modes, such as descriptive observations, focused observations, and selective observations.</p> <p>PIE graphic assessment method developed (cf. Cowan, 1988) qualitatively to assess “individuals’ and family members’ psychological commitment to the different roles in their lives” (Franklin & Jordan, 1995, p. 288). This qualitative assessment method may be extended to the context of education.</p> <p>The first use of the phrase <i>culturally responsive assessment</i> by Cuellar et al. (1983).</p> <p><i>Quantitative: Era of standards for quantitative assessment:</i></p> <p>APA, AERA, and NCME jointly published the third revision of the standards for educational and psychological assessment, which represented a shift toward a unitary concept in validity theory (APA, AERA, & NCME, 1985).</p> <p>The Joint Committee on Standards for Educational Evaluation published the Personnel Evaluation Standards in 1988, which included a total of 21 standards.</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
1990-1999		<p><i>Moment 5: Post-modern period of experimental ethnographic writing:</i></p> <p>Struggle to make sense of triple crises; new ways of composing ethnography emerged (e.g., auto-ethnography); concept of passive observer discarded; more action-, participatory-, and activist-oriented research emerged.</p> <p><i>Moment 6: Post experimental Inquiry:</i></p> <p>Writings connected to the needs of a free democratic society; experimental forms of qualitative writing published that blurred the boundaries between the social sciences and the humanities.</p>	<p><i>Institutionalization of mixed methods as a distinct methodological orientation:</i></p> <p>Beginning of conversations between quantitative and qualitative researchers.</p> <p>Publication of seminal works promoting mixed methods research as a separate research movement.</p> <p>Widespread publication of mixed methods research studies throughout the human sciences.</p> <p>Conceptualization that much research is inherently mixed.</p>	<p><i>Era of experimental ethnographic ways of assessment reporting:</i></p> <p>Emergence of action-, participatory-, and activist-oriented assessment.</p> <p>The social network map developed in 1990 for qualitatively assessing social support by taking into account both the structure and function of the individual's personal social network (Tracy & Whittaker, 1990). This social network mapping approach may be extended to the field of education to assess educational support.</p> <p>Repertory grids used "as a qualitative assessment tool for constructing personal meanings or constructs" (Franklin & Jordan, 1995, p. 286) and, therefore, are applicable to the field of education (see also Neimeyer, 1993). With respect to the field of education, repertory grids elicit a students' construction of some domain of experience by asking them to compare and to contrast representatives from that domain (e.g., classmates, teachers, administrators, family members) and then systematically to describe each of these representatives on their own repertory of dimensions of assessment, or personal constructs. The repertory grid can be administered formally orally (i.e., interview) or in writing—either offline or online (Neimeyer, 1993).</p> <p>Development of Stories/Narrative Assessment Procedure (SNAP), which is an innovative assessment procedure using stories, which records the development of the narrative (i.e., language and communication) skills of young deaf children (Starczewski & Lloyd, 1999; Strong et al., 1998).</p> <p><i>Era of post experimental inquiry:</i></p> <p>The Joint Committee on Standards for Educational Evaluation published the second edition of the Program Evaluation Standards in 1994.</p> <p>APA, AERA, and NCME jointly published the fourth revision of the standards for educational and psychological assessment, which emphasized that validity and reliability were functions of the interpretations of test scores for their intended uses and not of the test itself (APA, AERA, & NCME, 1999).</p> <p>New forms of assessment reporting blurred the boundaries between social sciences and humanities.</p>
2000-2009		<p><i>Moment 7: Methodologically contested present:</i></p> <p>Period of conflict, great tension, and retrenchment; growing body of literature on paradigms, approaches, and methods.</p> <p><i>Moment 8: Un-named</i></p> <p>Period of confronting the methodological ramifications of the evidence-based social movement.</p>	<p><i>Institutionalization of mixed methods as a distinct methodological orientation:</i></p> <p><i>Handbook of Mixed Methods Research</i> published (2003).</p> <p><i>Journal of Mixed Methods Research</i> launched (2007)</p> <p><i>International Journal of Multiple Research Approaches</i> launched (2007)</p>	<p><i>Qualitative: Era of diversity of assessment methods:</i></p> <p>Different assessment methods were pitted against each other (e.g., unimodality vs. multimodalities).</p> <p>Qualitative assessment occurred during the evidence-based social movement.</p> <p>Shift to explore multimodalities (Kress & Jewitt, 2003; Kress & Van Leeuwen, 2001) and multiliteracies (Cope & Kalantzis, 2000), which created openings for assessments that include the examination of various modalities and multimodal analyses.</p> <p>The Edmonton Narrative Norms Instrument (ENNI), developed in 2005, is an assessment tool for collecting language information from children aged 4–9 through storytelling. Pictures that portray a story are presented to a child, who then tells the story to the examiner (Schneider et al., 2005).</p> <p><i>Quantitative: Era of diversity of assessment methods:</i></p> <p>No Child Left Behind (NCLB) education reform led to the expansion of state-mandated standardized testing as a means of assessing school performance, wherein most students are tested each year of grade school (2001).</p> <p>Latent Dirichlet allocation (LDA) was developed, which is perhaps the most common topic model tool (i.e., a text-mining tool for discovery of hidden semantic structures in a body of text) (2002).</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
		<p><i>Moment 9: Fractured future:</i></p> <p>Methodologists form two opposing camps (i.e., “gold standard” of scientific research vs. socially, culturally, ethnically, and racially responsive, communitarian, justice-oriented research).</p>	<p>Special interest groups formed (e.g., American Educational Research Association).</p> <p>Emergence of dialectical pragmatism introduced as a metaparadigm (2009).</p>	<p>The Joint Committee on Standards for Educational Evaluation published the Student Evaluation Standards in 2003.</p> <p>Changes to the SAT to assess what students are learning in school (2005).</p> <p>Online Research and Comprehension Assessments (ORCA), which are performance-based measures of students’ ability to conduct online research and to write a short report of their results, were used with content stability issues, wherein target websites were subject to change during data collection (e.g., Henry, 2007). Therefore, “an assessment used at one time was not always comparable to the same assessment used at another time” (Leu et al., 2015, p. 42).</p> <p>The Joint Committee on Standards for Educational Evaluation published the second edition of the Personnel Evaluation Standards in 2008, which is based on knowledge about personnel evaluation gained from the professional literature and research/development since 1988. A total of 6 new standards were added to the original 21 standards of the first edition.</p>
2010-2014		<p><i>Moment 10: Methodological Innovation</i></p> <p>Utilization of innovative approaches to reflexivity and latest technology and computer-mediated communication.</p>	<p><i>Emergence of mixed methods research into young adulthood</i></p> <p>Second edition of <i>Handbook of Mixed Methods Research</i> published (2010).</p> <p>Mixed Methods International Research Association.</p> <p>Mixed Methods Research conferences held in multiple countries and continents.</p> <p>Webinars held.</p> <p>Dialectical pragmatism changed to dialectical pluralism (2011).</p> <p>Emergence of critical dialectical pluralism (2013).</p>	<p><i>Qualitative: Era of technological innovation in assessment:</i></p> <p>Use of the Narrative Assessment Protocol, which provides a direct assessment of children’s language abilities within a narrative context via assessing the following five aspects of language: sentence structure, phrase structure, modifiers, nouns, and verbs. It involves a real-time online scoring procedure (Justice et al., 2010).</p> <p>Multilingual Assessment Instrument for Narratives (MAIN) developed by the Working Group for Narrative and Discourse as a tool for the evaluation of the narrative abilities of bilingual children across a variety of languages and language combinations. The design of the MAIN allows for the elicitation of narratives in three modes: (1) story generation (telling), (2) retelling, and (3) telling after listening to a model story (Gagarina et al., 2012).</p> <p>Ladson-Billings’s (2014, 2017) concept of culturally relevant pedagogy, Paris and Alim’s (2014) extension of culturally responsive pedagogy (CRP) into culturally sustaining pedagogy (CSP), and Randall et al.’s (2022) work in justice-oriented antiracist assessment creates space for discussions of issues with assessments that marginalize people of color.</p> <p>Qualitative assessment of wiki-based learning processes emerged (Balderas et al., 2012).</p> <p><i>Quantitative: Era of technological innovation in assessment</i> (Mislevy, 2016, 2019):</p> <p>APA, AERA, and NCME jointly published the fifth revision of the standards for educational and psychological assessment (APA, AERA, & NCME, 2014).</p> <p>Continued development of ORCA (Leu et al., 2015).</p>
2015-Present	<p><i>Emergence of Equity Movements:</i></p> <p>Institute in Critical Quantitative, Computational, & Mixed Methodologies (ICQCM) (2020).</p>	<p><i>Moment 10: Methodological Innovation</i></p> <p>Publication of Sage textbook entitled “Conducting qualitative research of learning in online spaces” (Gerber et al., 2017).</p>	<p><i>Mixed Methods Research 2.0: Emergence of Integration</i></p> <p><i>Oxford Handbook of Multi method and Mixed Methods Research Inquiry</i> (2015)</p>	<p><i>Qualitative: Technological innovation in assessment, continued:</i></p> <p>Culturally sustaining pedagogies (Paris & Alim, 2017) and culturally relevant pedagogical assessment.</p> <p><i>Quantitative: Technological innovation in assessment, continued</i> (Mislevy, 2016, 2019):</p> <p>Continued development of ORCA (Leu et al., 2015).</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
<p>The <i>Me Too</i> (or #MeToo) (founded in 2006) movement began to spread virally as a hashtag on social media (2017).</p> <p>The Extinction Rebellion (XR) global environmental movement (2018).</p>			<p>The Comprehensive Literature Review Process framed as a Mixed Methodology (Onwuegbuzie & Frels, 2016).</p> <p><i>Emphasis on two forms of integration instead of mixing:</i></p> <p>(1) Emergence of “1 + 1 = 3” integration formula in 2015 (Fetters & Freshwater, 2015).</p>	<p>Ofqual (i.e., Office of Qualifications and Examinations Regulation)—which regulates qualifications, examinations, and assessments in England—created an algorithm to grade/mark GCSE (i.e., General Certificate of Secondary Education) and A-level (i.e., Advanced Level) examinations for determining university places in the United Kingdom because of students’ inability to take face-to-face examinations as a result of the COVID-19 lockdowns (2020). The algorithm, based on four distinct terms, factored in the school’s performances in each subject over the previous 3 years. The algorithm was abandoned soon after (August 13, 2020) when it was determined that, in England, Wales, and Northern Ireland, nearly 40% of the grades computed by the algorithm were lower than teachers’ assessments.</p> <p>In the United Kingdom, GCSE, AS, and A level examinations did not take place in 2021 due to the global COVID-19 pandemic. Instead, the students’ grades were assessed by teachers based only on what they had been taught, and not what they had missed.</p>
<p>The Black Lives Matter movement (founded 2013) gained further international attention during the global George Floyd protests (2020).</p>			<p>(2) Emergence of “1 + 1 = 1” integration formula in 2017 (Onwuegbuzie, 2017; Onwuegbuzie & Hitchcock, 2019a).</p> <p>Integration of multiple methods research and mixed methods research (Onwuegbuzie & Hitchcock, 2019b).</p> <p><i>The Routledge Reviewer’s Guide to Mixed Methods Analysis</i> (Onwuegbuzie & Johnson, 2021).</p> <p><i>The Routledge Handbook for Advancing Integration in Mixed Methods Research</i> (Hitchcock & Onwuegbuzie, 2022).</p>	<p>The National Council of Teachers of English published position statements, such as <i>Expanding Formative Assessment for Equity and Agency</i> (Beck et al., 2020); the <i>Definition of Literacy in a Digital Age</i> (Witte et al., 2019) that acknowledged that “today’s literacy demands have implication for how teachers plan, model, support, and assess student learning” (¶ 3); and <i>Literacy Assessment: Definitions, Principles, and Practices</i> (Yancey et al., 2018), which acknowledged that “literacy assessment is a social process not a technical activity...literacy assessment includes more than cognitive activities; it also includes a range of practices and perceptions, including beliefs about literacy, dispositions toward literacy, and self-efficacy regarding literacy” (¶ 10).</p>

Figure 1 continued

Period	Quantitative Research ^a	Qualitative Research ^b	Mixed Methods Research ^c	Qualitative Educational Assessments and Quantitative Educational Assessments in the Western World
			<p><i>Emergence of Equity Movements:</i></p> <p>Institute in Critical Quantitative, Computational, & Mixed Methodologies (ICQCM) (2020).</p> <p>Integrative Mixed methods Antiracist Groundwork for Investigating and Nurturing Equity (IMAGINE) movement (2021) (Abrams et al., 2021, 2022).</p>	

Notes

^a Johnson and Gray (2010); Teddlie and Johnson (2009).

^b Denzin and Lincoln (2011); Onwuegbuzie, Leech, and Collins (2010, 2011).

^c Johnson and Gray (2010); Johnson et al. (2007); Teddlie and Johnson (2009).

Adapted from Onwuegbuzie, A. J., Forzani, E., & Abrams, S. S. (2022). *History of quantitative, qualitative, mixed methods research, and educational assessment: A review*. Unpublished manuscript, University of Cambridge, Cambridge, England. Copyright 2022 by A. J. Onwuegbuzie, E. Forzani, and S. S. Abrams.

Purpose of this Article

For the remainder of this article, we invite researchers—regardless of their methodological experiences and orientations—to embrace a new research and evaluation meta-framework for transforming assessment to support more equitable and empowering research. This framework is driven by a research philosophy called *critical dialectical pluralism* (CDP; Onwuegbuzie et al., in press-a; Onwuegbuzie & Frels, 2013), which underscores the inclusion of participants and their voices at every stage of the research process, from conceptualization to dissemination to utilization. Adoption of CDP promotes research that honors participants as co-researchers, co-ideators, and co-decision makers, who, through their involvement in the research, become activists striving to identify problems and to effect change that they envision in local and/or global contexts. Both this new research and evaluation meta-framework and its philosophical underpinning (i.e., CDP), in turn, have at their core a movement that embraces an **I**ntegrative **M**ixed methods **A**ntiracist **G**roundwork for **I**nvestigating and **N**urturing **E**quity. It is to this movement, namely, IMAGINE (S. S. Abrams et al., 2021, 2022; Onwuegbuzie et al., in press-b)—which is devoted to re-imagining research to be more equitable—to which we now turn.

The IMAGINE Movement

Here, we embrace the first definition from the Cambridge Dictionary (2021b, ¶ 1) of the word, *imagine*—“to form or have a mental picture or idea of something”—as we envision what the future of (mixed methods) research could look like and how it could shift systems of power. After all, the ivory tower of academia and of related research is as ivory suggests: It is White, it is privileged, it is durable, and, historically, it has come at the cost of marginalizing others. Conventional research traditions support hierarchical knowledge (e.g., the researcher vs. the participant) and perpetuate “scientific racism” (Saini, 2019, p. 29), which dehumanizes people (typically people of color) and which, as Saini, quoting her discussion with

Professor Jonathan Marks, explained, “emerged ‘in the context of colonial political ideologies, of oppression and exploitation. It was a need to classify people, make them as homogeneous as possible.’ Grouping people made it easier to control them” (p. 29). Such grouping, exploitation, and othering has been part of experimentation-in-the-name-of-science throughout history—or what we call *assessment-in-the-name-of-science*. Furthermore, the human experimentation of the Holocaust and the resulting Nuremberg trials of 1945–1946 created an impetus and a policy (i.e., the Nuremberg Code) to ensure ethical research that, first and foremost, protects the rights of all research participants.

Although the Nuremberg Code and ethical research approaches exist at the forefront of contemporary investigations, there remains an underlying (and sometimes overt) othering, with particular respect to race, as noted previously (see also Figure 1). Furthermore, in a day-and-age when there is greater awareness of systemic racism, it is not enough to acknowledge that racism is embedded within these systems, including within research. Indeed, as declared by Onwuegbuzie (2021) in a recent Black Lives Matter special issue published in the *International Journal of Multiple Research Approaches*, “non-racism ≠ anti-racism, with the former representing a passive, and even perhaps a passive-aggressive, stance” (p. 106). (See also C. Abrams & S. S. Abrams, 2021.) The systemic racism that is perpetuated via research needs to be addressed immediately—not in the future, but *now!* There is a longstanding need for research, in general, and education research, in particular, to be re-imagined. And there is no more important way to address the systemic racism that is inherent in research than by *reforming assessment systems* that make questionable numerous past and present findings conducted on participants of color.

An immediate way that the reform of assessment systems in research can take place is via the IMAGINE movement. Developed in 2021, this movement, which represents the mixed methods research community, is dedicated to creating pathways for conducting research that

are integrative, and, even more importantly—by adopting an antiracist stance—that promote research that facilitates the cultivation of equity (S. S. Abrams et al., 2021, 2022; Onwuegbuzie et al., in press-b). This movement is dedicated to the reimagining of (mixed methods) research by dismantling hierarchies that create unnecessary and exclusionary dichotomies that privilege certain voices—specifically, voices belonging to the dominant majority group. The IMAGINE movement is built on the premise that the current status of research methods *must* shift to include antiracist methodological approaches that offer an equitable way to design, to conduct, to analyze, to interpret, and to disseminate research. Although these approaches should dismantle systems, the IMAGINE movement focuses less on destruction and more on reformation and reconstruction. One way to achieve such a focus on rebuilding is through the integration of culturally responsive assessment (cf. Cuellar et al., 1983; Logli, 2020) and antiracist assessment (Randall, 2022). Such equity-oriented assessment is holistic in nature and “calls for student involvement throughout the entire assessment process including the development of learning outcome statements, assessment tool selection/development process, data collection and interpretation, and use of results” (Montenegro & Jankowski, 2017, p. 10). In this way, culturally responsive assessment has an ethos similar to CDP: The traditional figure of authority—be it the adult, White researcher, or, in the classroom, the teacher—is not a central or a privileged figure. Assessment is designed and applied—and the power it garners is distributed—in meaningful and responsive ways.

Furthermore, culturally responsive assessment is embedded within the IMAGINE movement, which seeks to develop, to nurture, and to sustain research by future Methodologists of Color, including those from the Caribbean and other Global South countries and territories, within and across colleges and universities, K-12 classrooms, and home environments. After building on generative feedback received from mixed methods researchers representing the Caribbean (S. S. Abrams et al., 2022), Onwuegbuzie et al. (in press-b) have outlined how these under-

represented populations can be nurtured in their development as methodologists from primary school through tertiary education and beyond, with the goal of promoting equity within the (mixed methods) research community.

Critical Dialectical Pluralism

The IMAGINE movement is rooted in CDP and its pursuit of **s**ocial justice, **i**nclusion, **d**iversity, **e**quity, and **s**ocial responsibility (*SIDES* of CDP; Onwuegbuzie et al., in press-a). Broadly speaking, originally developed by Onwuegbuzie and Frels (2013), CDP 1.0 (circa 2013) is a mixed methods research-based philosophy that builds on dialectic pluralism (DP; circa 2009), which, in turn, has its roots in dialectical pragmatism (circa 2009) (cf. Figure 1). DP involves a belief in incorporating multiple epistemological perspectives within the same inquiry (Johnson, 2012, 2017; Johnson et al., 2014; Tucker et al., 2020). Like DP, CDP 1.0 represents both a process philosophy and a communication theory promoting both universalistic theoretical knowledge and local practical knowledge. Although one of the goals of DP researchers is to “‘give voice’ to those with the least power” (Johnson, 2012, p. 753) and to “reduce inequality” (Johnson, 2017, p. 165), this is not the major goal associated with this research philosophy—as evidenced by the lack of articulation, to date, as to how DP researchers give voice to those with the least power. In contrast, CDP 1.0 emphasizes procedural, process, and philosophical justice. CDP 2.0—hereafter referred to as CDP—builds on CDP 1.0, introduced by the authors of this present article (i.e., Onwuegbuzie et al., in press-a). This version of CDP has, what we refer to as, *the many SIDES of CDP* (i.e., **s**ocial justice, **i**nclusion, **d**iversity, **e**quity, and **s**ocial responsibility; Onwuegbuzie et al., in press-a), which represent five core elements that facilitate global justice (Al-Rodhan, 2009).

Most notably, CDP differs from critical-based, participatory-based, and transformative-emancipatory research philosophies. With respect to critical-based research philosophies, whereas CDP emphasizes mixed methods research approaches, critical-based research philosophies tend to promote mono-method research

studies—specifically, *quantitative research studies* (e.g., critical quantitative research [Baez, 2007; Teranishi, 2007]; quantitative criticalism [Stage, 2007]) and *qualitative research studies* (e.g., critical theory [Morrow & Brown, 1994]; critical race theory [Delgado & Stefancic, 2012]). Similarly, participatory-based research studies are predominantly qualitative in nature (cf. Hall et al., 2021), and do not consistently or fully honor the voices of participants and, relatedly, sometimes sustain hierarchical structures and stymie the participatory nature of participatory research (S. S. Abrams & Schaefer, 2022); thus, CDP has been a more apt frame for supporting participation at every stage of the inquiry and, as noted by Mertens (2007), “Methodologically, mixed methods are preferred for working toward increased social justice because they allow for the qualitative dialogue needed throughout the research cycle, as well as the collection of quantitative data as appropriate” (p. 224). Contrastingly, the transformative-empowerment philosophy involves the utilization of mixed methods research approaches that focus directly on the lives and experiences of underserved, under-represented, and marginalized individuals or groups, including, but not limited to, women; ethnic/racial/cultural minorities; certain religious groups, individuals with disabilities/exceptionalities; and members of lesbian, gay, bisexual, and transsexual (LGBT+) communities—yielding research that is participatory, antidiscriminatory, and emancipatory (e.g., Mertens, 2003, 2007, 2010; Mertens et al., 2010). However, although the goal of transformative-empowerment studies is to capture the voices of these individuals or groups, their voices are filtered through the voice of the researcher(s) (Onwuegbuzie & Frels, 2013). That is, studies emanating from this transformative-empowerment research philosophy—as well as from critical- and participatory-based research philosophies—privilege researchers over participants wherein the researcher(s) has ultimate power over the participant(s) in terms of decisions made at every stage of the

research process. These decisions include the research questions that should be addressed, the positionality of the researcher(s) and each of the participants, the aspects of each participant’s voice that are included and excluded, and the veracity with which each participant’s story is (re-)told (Onwuegbuzie et al., in press-a; Onwuegbuzie & Frels, 2013). That is, from the perspective of proponents of CDP researchers,

there are at least some occasions when using this paradigm does not go far enough in terms of giving voice to people who have been traditionally excluded, namely, those who represent disenfranchised and the least advantaged groups in society and who have the least power. (Onwuegbuzie & Frels, 2013, p. 13)

In contrast to researchers representing these aforementioned philosophies, and echoing the tenets of the IMAGINE movement, CDP researchers also emphasize the inclusion of participants and their voices at every stage of the research process as researchers work with participants as co-ideators, co-investigators, and, most importantly, co-decision makers. It is this attention to the many SIDES of CDP that led to the birth of Onwuegbuzie’s (2017) $1 + 1 = 1$ integration formula (see also Hitchcock & Onwuegbuzie, 2022; Natesan et al., 2019; Newman et al., 2015; Onwuegbuzie & Hitchcock, 2019a, 2019b, 2022; Onwuegbuzie et al., 2018; Onwuegbuzie & Johnson, 2021). With this formula, the researchers adopt an integrative, integrated, and integral way of thinking at all phases of the research process that promotes the full(er) integration of the following seven broad elements identified by Onwuegbuzie and Hitchcock (2022): (a) quantitative and qualitative research approaches, (b) mixed methods research and multimethod research approaches (i.e., meta-methods research study), (c) disciplines/fields (e.g., interdisciplinary, multidisciplinary, transdisciplinary, cross-disciplinary), (d) arts and sciences, (e) Global North and Global South researchers, (f) online and offline spaces, and, most importantly, (g) researchers and participants.

Although CDP is a relatively recent research philosophy—being in existence for 11 years at the time of writing—methodologists who adopt a CDP stance have conceptualized and implemented methods that attempt to reduce procedural and process injustice, including: (a) the development of sampling designs that are **t**ransparent, **r**igorous, **e**quitable, and **e**thical—what Corrigan and Onwuegbuzie (2023) refer to as being more *TREEful*—especially when sampling among/between phases/components; (b) the development of a meta-framework for optimal matching—wherein matching refers to the process of forming groups to make them as similar as possible in terms of extraneous or confounding factors (e.g., demographic variables [e.g., gender, race/ethnicity]; personality variables [e.g., resilience]; affective variables [e.g., motivation]) (Onwuegbuzie & Corrigan, 2021); (c) the development of focus group discussions that are designed and conducted by the research participants themselves, what the authors refer to as *critical dialectical pluralist focus group discussions* (Onwuegbuzie & Frels, 2015); (d) the development of an integrated mixed methods (IMM) approach for conducting design-based research (DBR)—what the authors refer to as *critical dialectical pluralist integrated mixed methods design-based research* (i.e., CDP-driven IMM-DBR; Onwuegbuzie, Forzani, et al., 2023); (e) the development of a meta-framework for conducting impact evaluations that facilitate the adoption of an integrative, integrated, and integral way of thinking—thereby facilitating the addressing of more complicated and complex evaluation problems—what the authors refer to as *critical dialectical pluralistic mixed methods-based impact evaluations* (Onwuegbuzie, Forzani, Hitchcock, et al., 2022); (f) the development of a method of collecting qualitative data via what is known as paired depth interviewing (also known as paired interviewing), which is defined as one researcher interviewing two people together for the purposes of obtaining information regarding how the pair perceives the same event, experience, or phenomenon (Wilson et al., 2016); (g) the development of a method to

transform bibliometric studies to *mixed methods bibliometric studies* that involve the integration of bibliometric studies—which help researchers to determine the degree of development of various disciplines—and CDP-based qualitative research approaches (Onwuegbuzie et al., 2018); and (h) the development of a qualitative research process that ended in one or more of the participants (co-) presenting the findings at professional meetings (S. S. Abrams et al., 2017; Gerber et al., 2014; Onwuegbuzie et al., 2014; Schaefer et al., 2018).

Assessment Systems Re-IMAGINEd

The first step in re-IMAGINE-ing is to revise and to expand the definition of assessment from the macro-definition (i.e., general definition) of “the act of judging or deciding the amount, value, quality, or importance of something, or the judgment or decision that is made” (Cambridge Dictionary, 2021a, ¶ 1) to a micro-definition (i.e., specific definition)—inspired by Onwuegbuzie and S. S. Abrams (2021, in press)—that has been established by IMAGINE, and that is displayed in Figure 2. Adopting this definition allows us to envision and to create assessments that promote social justice, inclusion, diversity, equity, and social responsibility (i.e., SIDES), such as using a Critical Assessment Practices, or CAPS (Forzani, Dobbs et al., 2024) approach in research, which we will introduce in the next section.

Additionally, with research, in general, and when re-IMAGINE-ing assessment, in particular, acknowledging positionality is crucial, especially in light of culturally responsive and antiracist approaches. Each of us has a background in education—in the classroom as a K-12 educator and as a university professor—and each of us has had experience researching teaching and learning in various capacities. We all share a similar desire to see greater equity in research and in practice. We also abhor traditional assessment practices that continue to denigrate people of color and to undermine culturally responsive and antiracist research and teaching practices.

Figure 2

A Micro-Definition of Assessment

In the context of studying human beings, *assessment* is the process of documenting, analyzing, and interpreting empirical data to provide evidence regarding an observable human entity that emerges cognitively, affectively, physically, metaphysically, and/or spiritually, and which is internal or external to the person or group being observed. In its original form, the entity being assessed represents multimodal information that resides in nondigital spaces and/or digital spaces (e.g., spaces that are online and virtual and that exist via a wired or wireless connection) that are represented cognitively, affectively, physically, metaphysically, or spiritually. Once observed, this information is socially constructed, co-constructed, and co-produced by both the assessor(s) and those being assessed as representing qualitative data, quantitative data, or multidata (e.g., data that are neither exclusively qualitative nor quantitative; e.g., spirituality) that are generated continuously or discretely; consciously, subconsciously, or unconsciously; internally or externally; verbally or nonverbally; deliberately or accidentally; once or multiple times; individually, competitively, cooperatively, and/or collaboratively; and reflecting some level of learned or innate characteristic behavior. Further, via one or more of the human senses (e.g., vision, hearing, taste, smell, touch, intuition, proprioception [i.e., the perception of body position], temporal perception [i.e., the sense of the passage of time], interoception [i.e., sensations coming from within organs], nociception [i.e., physiological pain], thermoception [i.e., ability to feel hot and cold], equilibrioception [i.e., perception of balance or acceleration], radiation senses [e.g., sense of color, sense of moods associated with color, sense of temperature], feeling senses [e.g., sensitivity to gravity, air and wind pressure, and motion], chemical senses [e.g., hormonal sense, such as pheromones, hunger for food, water or air], mental senses [e.g., pain, external and internal, mental or spiritual distress, sense of self, including friendship, companionship and power, psychic capacity]), the assessment is undertaken on one or more persons or groups. This assessment can occur objectively, subjectively, and/or intersubjectively; formally or informally; systematically or unsystematically; while lying somewhere on the structured—unstructured continuum. These methods represent approaches that are *intradisciplinary* (i.e., involving a single discipline), *cross-disciplinary* (i.e., involving one or more disciplines/fields that are viewed from the perspective of another discipline/field), *multidisciplinary* (i.e., involving assessors that represent different disciplines/fields, with each assessor drawing on their disciplinary knowledge), *interdisciplinary* (i.e., involving the integration of methods of assessment from different disciplines), or *transdisciplinary* (i.e., involving a unity of assessment frameworks that are created in a way that is beyond the disciplinary perspectives). The resulting assessment is non-static and non-immutable (i.e., changing with the person[s] being assessed and time), situated in its use, and reliant on the purpose for the assessment and its intended consequences in the very process of providing information (i.e., data) in general and evidence in particular. The final communicative assessment product either can stand alone or can be combined or integrated with other assessment products in order to (a) repeat the evidence (i.e., provide the exact repetition of an assessment, using the same assessment process, and under the same conditions; *repetition*), (b) replicate the evidence (i.e., recreate the same assessment tool and use it to undertake exactly the same assessment; *replication*), (c) reproduce the evidence (i.e., implement the same general assessment, in a similar setting, with a newly created appropriate assessment tool; *reproducibility*), (d) corroborate the evidence (i.e., provide evidence in support of a hypothesis, hunch, or a result, but using a different assessment approach from the one used originally; *corroboration*), (e) compare the evidence (i.e., *triangulation*); (f) capture the underlying evidence (i.e., *complementarity*; cf. Greene et al., 1989, p. 258); (g) accentuate the underlying evidence (i.e., *accentuation*); (h) substitute the underlying evidence (i.e., *substitution*); (i) expose inconsistencies and contradictions with the underlying evidence (i.e., *initiation*; Greene et al., 1989, p. 260); (j) broaden the scope of the evidence (i.e., *expansion*; Greene et al., 1989, p. 260); (k) moderate the underlying evidence (i.e., *moderation* or *regulation*); (l) mediate the underlying evidence (i.e., *mediation*; e.g., filter); and/or (m) create new directions based on additional evidence (i.e., *development*; Greene et al., 1989, p. 260).

Adapted from Onwuegbuzie, A. J., Forzani, E., & Abrams, S. S. (2022). *History of quantitative, qualitative, mixed methods research, and educational assessment: A review*. Unpublished manuscript, University of Cambridge, Cambridge, England.
Copyright 2022 by A. J. Onwuegbuzie, E. Forzani, and S. S. Abrams.

Using a Critical Assessment Practices Approach in Research

Assessment used in research can include a wide range of approaches and types, from conducting surveys to evaluating programs, to developing assessment tools, to using informal assessment as part of design-based research (cf. Onwuegbuzie, Forzani, et al., 2023), to administering formal assessments to evaluate participants before, during, and after interventions. In other words, assessments often are used by researchers as tools to evaluate participants, groups, interventions, treatments, and programs. In this sense, assessment is done *to*, rather than *with*, participants. In research, then, assessors often focus on *outcomes* (i.e., assessment *results*) without focusing on the *process* of assessment and without considering how what they have learned from assessment informs their research.

Unfortunately, such traditional approaches to the use of assessment in research, therefore, often assume a power differential between assessor and assessed, which can result in negative effects on research participants. In particular, assessment historically has negatively affected people of color and other minoritized populations to a greater degree, compared to more majority populations (Lee, 2016). Further, as critical dialectical pluralists, we reject the existence of false dichotomies in research, such as objective versus subjective, emic versus etic, constructivism versus realism, human science versus natural science, universals versus particulars, relativism versus absolutism, quantitative research versus qualitative research, mixed methods research versus multiple methods research, and, most importantly, researcher(s) versus participant(s)—to name just a few false dichotomies. In the context of assessment, in particular, we reject the false dichotomy between assessor(s) and assessed. Rather, we maintain a synechist (i.e., anti-dualistic) stance by viewing assessment systems as situating a series of continua instead of dichotomies. For example, with regard to the distinctions between assessments in digital and nondigital spaces, we dismiss any ontological prioritization between those spaces; instead, akin

to the porous and blurred boundaries of meaning making within and across digital and nondigital spaces (Burnett & Merchant, 2014), we contend that the border between digital assessment and nondigital assessment is porous and blurred. For instance, assessment of the competence of the pilot of a fighter jet drone (i.e., an unmanned combat aerial vehicle; a profession that takes place in a nondigital location) can occur in both a nondigital space (i.e., assessing the pilot's skills during an actual flight) and a digital space (i.e., via a simulator); similarly, assessment of the competence of a military pilot—a profession that takes place nondigitally (e.g., via fighter jets) and digitally (i.e., via drones, use of digital equipment)—can take place in both spaces. Extending to the classroom, we see assessment existing across digital and nondigital domains, including, but not limited to, how tests, scores, interaction, and feedback are generated, delivered, and communicated. Students who engage in a cooperative assessment (S. S. Abrams, 2017, 2021a, 2021b, 2021c) might use digital tools (e.g., calculators, online sites) or they might enter their answers in a digital forum, but they also have the nondigital interaction, as well as any work completed on scrap paper. Add a videoconference platform to the mix (Abrams, 2021b), and the dichotomy between digital and nondigital becomes even more problematic and problematized. Thus, we think it is important, as researchers, critically to consider our approach to assessment in research and how such an approach, instead, might align with the principles of CDP and the ethos of the IMAGINE movement.

Therefore, here, we take up a view of assessment as research (Huot, 2002), or as an inquiry process in which information is collected to understand a particular question or a set of questions about students. Also, we apply CDP to assessment in research, or what Forzani, Dobbs, et al. (2024) developed and refer to as a Critical Assessment Practices (CAPS) approach. In this article, we focus specifically on those assessment processes that are taken up in educational research, as opposed to those that are taken up in classrooms. For a practice-based discussion of a CAPS approach in K-12 education, see Forzani, Dobbs, et al. (2024).

Participants as Empowered Partners: Assessing *with* and *for* Participants to Inform Research

Drawing on CDP, a key feature of a CAPS approach is the equity-oriented collaboration between researcher(s) and participant(s), or between assessor(s) and assessed, when assessments are used in research. Whereas typical assessments, and especially those used by researchers when collecting data, assume a large power differential between assessor and assessed, a critical approach to assessment in research aims to break down this power differential as researchers and participants co-construct research and assessment goals, assessments, analyses, and interpretations together from the beginning to the end of the research process. Thus, in a CAPS approach, as in CDP, researchers, participants, and other partners and collaborators develop a reciprocal relationship wherein the goals of different partners are discussed and revisited throughout the research process.

In the remainder of this article, we extend the CAPS approach, recognizing that at the heart of a CAPS approach is the role of the *participant-as-assessor* and various reconfigurations of power structures so that assessment not only is developed or co-developed by the participants, but also the approach to assessment honors participants' "funds of knowledge" (Moll et al., 1992), or "historically accumulated and culturally developed bodies of knowledge and skills" (p. 133). This supports adults (in this case, participants and researchers) to "know the [participant] as a 'whole' person, taking into account or having knowledge about the multiple spheres of activity within which the [participant] is enmeshed" (pp. 133-134).

When we assess *with* participants, as in a CAPS approach, participants' voices are central, as is their role as autonomous individuals who make decisions about assessment, determining what the assessment is for and how it can best serve them. This is in contrast to much current research, in which researchers determine the purpose of assessment and how it might serve *them*. Instead, in line with CDP, in a CAPS

approach, participants make decisions at every stage of the assessment process, alongside researchers and other collaborators. This is more likely to result in assessment that is *with* and *for* participants because not only do participants have a voice in this process, but also their voice is centered (see Forzani et al., 2020). In the following section, we outline briefly four key practices in such a CAPS approach.

1. Relationships First: Participants and Researchers Develop Dynamic Understandings of Themselves and of One Another

CDP "assumes a communitarian view of power that is represented by reciprocity between researcher(s) and the participant(s)—a relationship not of domination, but of intimacy and vulnerability" (Onwuegbuzie & Frels, 2013, p. 14). Because traditional assessment assumes a power imbalance between assessor and assessed, we argue here that the first thing researchers and participants need to do is to work actively to develop a reciprocal relationship whereby there is shared power. Without this, assessments will not be especially valid, credible, or productive. However, if researchers and participants work together to develop shared power in various degrees of partnership (e.g., S. S. Abrams et al., 2020b), then they can understand better one another's needs and work together towards shared goals.

Before engaging in building a reciprocal relationship, however, it is useful for each person to examine their own identities and positionalities to understand better the ways in which they might approach relationship building with others. In other words, before we actively can seek to break down power imbalances, we first need to be aware of such imbalances. Moreover, in order to build trust, intimacy, and vulnerability, participants and researchers need first to know and to understand themselves and one another and to build a relationship together. The approaches that follow outline three ways to begin understanding self and others.

Reflexivity: Examining our Own Positionalities

CDP calls on researchers to engage in reflexivity by examining “their own biases...and those of their participants” (Onwuegbuzie & Frels, 2013, p. 16) before and throughout research, in order to understand the ways in which our identities might impact our interpretation of findings. Applied to a CAPS approach, participants and researchers each explore their positionality, or their “perspective shaped by the researcher’s unique mix of race, class, gender, nationality, sexuality, and other identities” (Mullings, 1999, p. 337). In doing this work, both participants and researchers might explore their “enduring social identities [that] confer a status that enables or limits the exercise of power” (Frost & Holt, 2014, p. 90). Such a practice of exploring our positionalities is important because it helps us to develop an awareness of how different aspects of our identities might influence our interpretations of different pieces of information garnered from assessments—including results—at different stages of the assessment process (Collins et al., 2013; Frels & Onwuegbuzie, 2012; Leech & Onwuegbuzie, 2008; Onwuegbuzie et al., 2008). It also helps us to develop an awareness of how our own identities might influence the ways in which we interact with others; this means that engaging in reflexivity is an important practice prior to developing a relationship with others.

Reflexivity: Learning to Self-Assess

Similar to Schön (1983), who advocated for teachers to engage in reflection-in-action while teaching, as well as in retrospect through reflection-on-action, we contend that an IMAGINE framework for mixed methods research involves reflexive activity throughout and after the research process. However, looking to the SIDES of CDP 2.0 (i.e., **s**ocial justice, **i**nclusion, **d**iversity, **e**quity, and **s**ocial responsibility) to situate the research process and the reflexivity therein, we argue that the IMAGINE framework opens up the seemingly insulated and guarded realm of assessment. This is not to say that constructs of traditional assessments are hermetically sealed and solidified (even if they might feel that way);

rather, we envision the SIDES of possibility—that is, the expansion of what assessment is and what it could be when participants engage in levels of partnership in the research, in general (e.g., S. S. Abrams et al., 2020a, 2020b; Schaefer et al., 2020, 2021), and in mixed methods research, in particular (Onwuegbuzie, 2020).

Reflexivity is an important part of exploring positionality and bias (see, for e.g., Collins et al., 2013; Frost, 2016). For assessment that is driven by the participant-as-researcher, reflexivity is key for participants to develop and to hone their understandings of assessment and of self (see S. S. Abrams, 2016, 2018, 2021a). In the classroom, self- and peer-assessment can involve verbal and nonverbal communication (S. S. Abrams, 2017), and mixed methods research featuring such reflexivity can include verbal and nonverbal analyses (Onwuegbuzie & S. S. Abrams, 2021, in press). Furthermore, methods, such as video-elicited reflection (i.e., structured and unstructured cued retrospective reporting), similar to stimulated recall (S. S. Abrams, 2009, 2015, 2018; Lyle, 2003), can be useful for participants to view and to perceive retrospectively their engagement in a particular activity, thereby supporting the development of reflexivity. Likewise, participant-to-participant peer feedback vis-à-vis formal and informal cooperative activities, which can be generated in-action (i.e., while an activity is happening), is central to supporting participant-driven approaches that redefine the boundaries of assessment. Such boundary-pushing is evident in cooperative assessments (S. S. Abrams, 2016, 2017, 2018, 2021a, 2021b, 2021c, 2022), which reduce anxiety, support perseverance and problem solving, and honor participant agency, all of which work to redefine how assessments are formed and situated in the context of the educational experience. In other words, through cooperative assessments, as well as self- and peer-assessments, there are opportunities for participants—who otherwise are the subjects of scrutinizing assessments—to be the originators, the owners, and the developers of assessments that they deem as fair, as relevant, and as meaningful.

Developing a Reciprocal Relationship

Examining our own positionalities allows us to be prepared better for listening to others because we gain some understanding of how our own identities affect our interpretations of others' beliefs and interpretations of the world. Understanding our own identities also allows us to understand what motivates us to engage in particular practices; such understanding also can be used to consider what motivates others. Such an understanding then can form the basis of a relationship of shared power because both participants and researchers understand that, although they might have different perspectives or even different purposes, they also might have some common ground. Identifying these areas of common ground then allows both participants and researchers to begin to develop shared goals, a process that we discuss in the next section. Such work involves deliberately making the time and space to engage in ongoing conversations to learn about one another.

2. Shared Goals Among All Partners Developing Shared Outcomes

Often, researchers, participants, and other partners have somewhat different goals for assessment. Discussing these goals and how the assessment process can serve the needs of different kinds of partners simultaneously can facilitate the co-construction of shared goals, assessment design and materials, interpretations, and application to future contexts. In particular, we argue for shared goals that promote greater equity both among partners and when applied to the broader communities that the research will affect. Given that conceptions of validity in any particular assessment represent the broader social values of the assessment developers (Messick, 1995)—and thus can represent the social values of a single, dominant group developing an assessment—considering the ways in which the constructs we are measuring represent participants' cultures and backgrounds can help ensure that we are measuring constructs valuable to the people using them (Randall et al., 2022). This means that if, for example, researchers and participants are using a reading comprehension

assessment in their research, they might interrogate existing constructions of reading comprehension, considering what constructions count and for whom, and developing their own, more culturally valid construction appropriate to the context (see, for e.g., Forzani et al., 2024).

Co-Constructing, as Equal Partners, Shared Ways of Knowing and Coming to Know

Assessment is a method of better understanding a person or group of people, including understanding people's characteristics, thoughts, beliefs, and emotions, or what they know and can do (Mislevy, 2019). Because the purpose of educational assessment is to understand people, it is important for anyone conducting and interpreting assessments to remember that people are unique and to consider (a) how different people, and different groups of people, approach the world, and (b) how they think about knowledge and coming to know.

CDP calls on researchers to engage in *culturally progressive* (Onwuegbuzie & Frels, 2013, 2016), *culturally responsive*, and *culturally engaged* (Onwuegbuzie et al., in press-b) research that strives towards developing cultural awareness and understanding of different partners involved in the research process. Here, we argue that the process of seeking such understanding should be reciprocal, with participants and researchers each working to understand one another's cultural beliefs and practices. However, we also think it is important that participants' voices, in particular, are honored and elevated given the inherent and historical power imbalance associated with assessment.

Developing cultural awareness and understanding is especially crucial for engaging in assessment processes, including developing assessment purposes and procedures, implementing assessments, and interpreting results. According to Onwuegbuzie and Frels (2013, 2016), culturally progressive research can include developing (a) cultural awareness of beliefs; (b) cultural knowledge; and (c) cultural skills. Because assessment measures people's abilities and understandings, test-

takers' cultural beliefs, knowledge, and skills will influence the ways in which participants engage in an assessment. Likewise, researchers' cultural beliefs, knowledge, and skills will influence the ways in which they interpret assessments. Different cultural beliefs also can lead two different people, or two different groups of people, to draw two different interpretations of assessment results. Thus, taking the time to understand one another's funds of knowledge (Moll et al., 1992), cultural resources, and socially situated ways of being (e.g., beliefs, knowledge, and skills; Gee, 1989, 1996, 2000, 2011) is an important step for all partners to take prior to engaging in assessment practices, including interpretation. Moreover, all partners can debrief assessment processes together to understand better how processes, and their roles in different processes, might have affected outcomes. For example, after implementing assessment and before examining findings, participants and researchers deliberately can engage in structured debriefing sessions to inform their interpretations of findings (Collins et al., 2012; Frels & Onwuegbuzie, 2012; Onwuegbuzie et al., 2008; Weinbaum & Onwuegbuzie, 2016). This makes it more likely that both researchers and participants will understand better how their own and others' interpretations are influenced by their own lenses and cultural resources and that all partners will come to more accurate and comprehensive shared understandings.

3. All Partners are Empowered to Participate as Equal Partners in Decision Making at Every Stage of the Research Process

Following a CDP lens, researchers adopt a "researcher-facilitator role that empowers participants to assume the role of participant-researchers" who collaborate throughout research, including by "perform[ing] or present[ing]...findings" (Onwuegbuzie & Frels, 2013, p. 15). When assessors adopt a *facilitator* role, students are empowered to be partners in all aspects of assessment. As researchers and facilitators work alongside one another, they make decisions together at all stages of the assessment process, from planning and development, to

implementation, to interpretation, to application, and to dissemination. This makes it more likely that assessment will be conducted and used in ways that serve participants, or that will work *with* and *for* participants. For example, after co-developing constructs for what will be assessed, participants and researchers can work together to develop assessments and assessment scoring criteria (e.g., rubrics), and to be equal partners in interpreting, disseminating, and deciding how to use findings.

4. Developing Dynamic Understandings of Participants

Traditionally in research, assessments have been conducted at few time points, using only one or two contexts, and using only a single method of data collection (Pearson et al., 2014). Such an approach assumes that learners have static identities that do not vary by time, situation, or context. However, research suggests that people's identities and abilities do vary by text, activity, and context (see, for e.g., RAND Reading Study Group, 2002). Thus, when we use assessment in research, multiple assessment methods, mixed methods, and multiple timepoints are needed for developing comprehensive understandings of participants' practices and how such practices differ across contexts (Pearson et al., 2014). Rather than using measures that utilize only one method, participants and researchers can develop multi-method- (i.e., multiple methods-) and mixed methods-based assessments. Findings then can be integrated to form a more cohesive and nuanced understanding of a participant as a whole person across multiple methods, timepoints, and contexts. This means that rather than drawing conclusions about research data from a single assessment, or a single type of assessment, our conclusions can be more robust and comprehensive and, thus, more representative of those being assessed.

Findings also can be more robust and accurate because participants engage in the construction of assessment goals and tools, as well as in the interpretation of the findings. Rather than

researchers drawing conclusions about a student from assessment data, then, participants have a voice in how their assessments are interpreted. This both empowers participants and leads to more accurate interpretations of what participants know and can do, leading to more valid interpretations of the answers to our research questions.

Summary and Conclusions

Indubitably, assessment is the most important component of *all* research studies—whether they represent quantitative, qualitative, or mixed methods research studies—because it is impossible to arrive at any finding without some form of assessment taking place. Unfortunately, as described earlier (see Figure 1), throughout history—especially since the turn of the 20th century—assessments often have been misused and abused as a result of application of a narrow view of assessment—a view that maintains or even widens the power differential between the assessor(s) and assessed, the researcher(s) and participant(s), and which exacerbates racial and ethnic bias in reporting and interpreting bias.

With this in mind, we would like to think that this article has made a contribution to the literature in the following four ways. First, we have identified a research philosophy, namely, CDP, that lends itself to redressing the power imbalance inherent in virtually all assessment systems. This stems from the fact that CDP (a) is committed to research that promotes and that sustains an egalitarian society, (b) adopts a participant-centered approach to assessment, (c) changes the role of (mixed methods) researchers/assessors to (mixed methods) research-facilitators/assessment-facilitators and changes the role of participant/assessed to participant-researcher/participant-(self-) assessor, (d) leads to the research-facilitator(s)/assessment-facilitator(s) serving as a “cultural broker” (Onwuegbuzie & Frels, 2013, p. 15) between the participant-researchers/participant-(self-)assessors and the assessment system that has power over them, (e) leads to the research-

facilitators/assessment-facilitators adopting a communitarian view of power and problematizes powerlessness, and (f) motivates the research-facilitator(s)/assessment-facilitator(s) to be committed to promoting social justice (i.e., equitably allocate power), distributive justice (i.e., equitably allocate resources), retributive justice (i.e., notion of merit), restorative justice (i.e., via conflict resolution), compensatory justice (i.e., for violations of rights), and, above all, global justice (i.e., comprising the following 8 minimum criteria: dialogue, effective and representative multilateral institutions, representative decision-making structures, fair treatment, empathy, accountability, transparency, and adherence to international law [Al-Rodhan, 2009]).

Second, we have applied a research movement that we have co-created recently, namely, the IMAGINE movement (S. S. Abrams et al., 2021; Onwuegbuzie et al., in press-b)—which is devoted to using CDP for the pursuit of social justice, inclusion, diversity, equity, and social responsibility—to assessment in research. Among numerous research elements (e.g., sampling designs used in mixed methods research; Corrigan & Onwuegbuzie, 2023; Onwuegbuzie & Corrigan, 2021), these five SIDES of CDP are all affected by past and present assessment systems. A major goal of this movement is to demonstrate how assessment systems can be re-IMAGINED to promote an antiracist and equitable approach to the development and use of assessments in research. Simply put, the IMAGINE movement is committed to playing a role in positively changing the assessment culture for both researchers and practitioners (e.g., primary-, secondary-, and tertiary-level teachers; test developers).

Third, we significantly have expanded the narrow definition of assessments found in textbooks and dictionaries. Specifically, using the tenets of CDP, we provided what we refer to as a micro-definition. This micro-definition creates spaces for researchers, in general, and mixed methods researchers, in particular, to develop multi-method- and mixed methods-based assessments that facilitate a more cohesive and nuanced understanding of a participant as

a whole person across multiple timepoints and contexts.

Fourth, we have described how a Critical Assessment Practices (CAPS) approach (Forzani, Dobbs, et al., 2024) in research can help to re-IMAGINE assessment systems. Inherent in this approach is the collaboration between and/or among researcher(s) and participant(s), or between and/or among assessor(s) and assessed, when assessments are used in research, wherein a reciprocal relationship is nurtured between and/or among these groups. What comes to the fore is the importance of participation *in*, and ownership *of*, assessment (S. S. Abrams, 2017, 2021a, 2021b, 2021c, 2022), which underscores the re-envisioning of how and why assessments can be developed, adapted, and applied to understand better integrated participant meaning making. Further, this stance, combined with the CAPS approach, which helps to facilitate the role of the participant-as-assessor, should enable partners to use the assessment system to refine interpretations vis-à-vis methods, such as repetition, replication, reproducibility, corroboration, triangulation, complementarity, accentuation, substitution, initiation, expansion, moderation/regulation, mediation, and/or development.

We realize that our call for researchers to design and to implement assessment systems that are re-IMAGINED adds a layer of complexity and complication to the research process—whether it be a quantitative research study, a qualitative research study, or a mixed methods research study. However, we believe that this is offset by the fact that such assessment systems facilitate a greater commitment to **s**ocial justice, **i**nclusion, **d**iversity, **e**quity, and **s**ocial responsibility—reflecting the *many SIDES of CDP*—than has been the case with previous assessment systems (see Figure 1). Therefore, we hope that this article represents an important step towards utilizing assessment systems in research studies that create a space for the empowerment of research participants—consistent with the principles of CDP and the ethos of the IMAGINE movement.

Anthony J. Onwuegbuzie, PhD is a Distinguished Visiting Professor at the University of Johannesburg and a Professor Extraordinarius at the University of South Africa. His research areas primarily involve social and behavioral science topics, including disadvantaged and under-served populations such as minorities and juvenile delinquents. Additionally, he writes extensively on qualitative, quantitative, and mixed methodological topics. With an h-index of 114, he has secured the publication of more than 550 works, including more than 400 journal articles, 55 book chapters, and 6 books. He has delivered more than 1,000 presentations, 250 methodological workshops, and 70 keynote addresses worldwide. He is former editor of *Educational Researcher*. Currently, he is editor-in-chief of both the *International Journal of Multiple Research Approaches* and the *Journal of Mixed Methods Studies*. He is past President of the Mixed Methods International Research Association. Professor Onwuegbuzie was a recipient of the 2024 National Research Foundation (NRF) top honors awards in the science and research field—receiving an “A” rating.

Correspondence concerning this article should be addressed to Anthony J. Onwuegbuzie, Faculty of Education, B-Ring 433, Kingsway Campus, Auckland Park, University of Johannesburg, South Africa, or

Email: tonyonwuegbuzie@aol.com

Elena Forzani, PhD is an assistant professor in Literacy Education at Boston University’s Wheelock College of Education & Human Development, where she teaches undergraduate and graduate courses in literacy instruction and assessment, with a focus on technology. Her research centers on using multiple and diverse methods and perspectives to understand and support digital literacies practices across the elementary and secondary levels. In particular, her work investigates the cognitive, metacognitive, and motivational dimensions of online reading comprehension, with special attention to the critical evaluation of online information. Through this work, Dr. Forzani seeks to inform the design and implementation

of assessment and instruction that supports different kinds of learners as they engage in reading and inquiry in complex, multiple-text online environments. Following her doctoral work in educational psychology at the University of Connecticut, Elena worked at Boston College as the assistant research director for PIRLS and ePIRLS, international print and digital reading assessments. Prior to earning her doctorate, Elena taught first grade and high school English and Reading, and she served as a high school literacy specialist.

Email: eforzani@bu.edu

Sandra Schamroth Abrams, PhD, is Professor of Adolescent Education and Director of the Ph.D. program in Curriculum and Instruction at St. John's University, N.Y. Abrams's research of videogaming, literacies, and pedagogies explores layered meaning making across digital and nondigital spaces. Awarded the St. John's University Outstanding Achievement Medal, Abrams has received research and publication awards from the United States Distance Learning Association and from the Association for Educational Communications and Technology. Abrams is author/editor of more than 80 works, including 8 books and 6 special issues. Abrams serves on multiple editorial boards, and she is an Associate Editor for the International Journal of Multiple Research Approaches and a founding co-editor of the Gaming Ecologies and Pedagogies book series (Brill). Along with Tony Onwuegbuzie, Abrams is co-author of *An integrated mixed methods approach to nonverbal communication data: A practical guide to collection and analysis in online and offline spaces* (Routledge).

Email: abrams.sandra@gmail.com

References

- Abrams, C., & Abrams, S. S. (2021). Dialogue is anything but silent: A foreword to the Black Lives Matter Special Issue. *International Journal of Multiple Research Approaches*, 13(1), 9-18. <https://doi.org/10.29034/ijmra.v13n1foreword>
- Abrams, S. S. (2009). A gaming frame of mind: Digital contexts and academic implications. *Educational Media International*, 46(4), 335-347. <https://doi.org/10.1080/09523980903387480>
- Abrams, S. S. (2015). *Integrating virtual and traditional learning in 6-12 classrooms: A layered literacies approach to multimodal meaning making*. Routledge.
- Abrams, S. S. (2016, October 26-28). *Socially responsible learning: Collaborative competition and reflective communication in public high school math classes* [Paper presentation]. Northeastern Educational Research Association, Trumbull, CT, United States.
- Abrams, S. S. (2017). Cooperative competition, reflective communication, and social awareness in public high school math classes. In Y. Baek (Ed.), *Game-based learning: Theory, strategies and performance outcomes* (pp. 357-370). Nova.
- Abrams, S. S. (2018, April). *Collaborative testing and game-based learning in high school mathematics classes* [Roundtable paper]. American Educational Research Association, New York, NY, United States.
- Abrams, S. S. (2021a). *Cooperative understandings of assessments, of peers, and of self*. [Unpublished manuscript]. Department of Curriculum and Instruction, St. John's University.
- Abrams, S. S. (2021b). Reimagining numeracies: Empowered game-informed meaning making in and beyond the pandemic era. *Language & Literacy*, 23(2), 16-31. <https://doi.org/10.20360/langandlit29551>
- Abrams, S. S. (2021c). Game-informed cooperative assessments and socially responsible learning in public school math classes. In M. Harvey & R. Marlatt (Eds.), *Esports research and its integration in education* (pp. 168-184). IGI Global. <https://doi.org/10.4018/978-1-7998-7069-2.ch010>
- Abrams, S. S. (2022). Game-informed meaning making in U. S. math class: Cooperative competition and students' literacies and numeracies. *L1: Educational Studies in Language and Literature*, 22(2), 1-25. <https://doi.org/10.21248/l1esll.2022.22.2.368>

- Abrams, S. S., Onwuegbuzie, A. J., Forzani, E., Natesan Batley, P., Schaefer, M. B., Abrams, C., Abrams, M., & Kurtis, M. (2022, March 25). *A meta-framework to Re-IMAGINE the increase of Black methodologists in education research: Integrative mixed methods antiracist groundwork for investigating and nurturing equity* [Workshop presentation]. Mixed Methods International Research Association—Caribbean Chapter, in collaboration with the School for Graduate Studies and Research, University of the West Indies.
- Abrams, S. S., Onwuegbuzie, A. J., Shannon-Baker, P., Hitchcock, J., Hines, M., Harris, A., Fletcher, E. C., Tan, T. X., Holmes, A. E., Bell, D., Decker-Woodrow, L. E., Barfield, D., & Sealy-Ruiz, Y. (2021, April 21-26). *A call to re-IMAGINE mixed methods research in the Black Lives Matter era* [Symposium]. American Educational Research Association, San Diego, CA, United States.
- Abrams, S. S., & Schaefer, M. B. (2022). Reimagining literacies through child-parent collaborations. *Journal of Adolescent & Adult Literacy*, 66(3), 142-145. <https://doi.org/10.1002/jaal.1270>
- Abrams, S. S., Schaefer, M. B., & Ness, D. (Eds.) (2020a). *Child parent research reimagined*. Brill.
- Abrams, S. S., Schaefer, M. B., & Ness, D. (2020b). Reimagining child-parent research. In S. S. Abrams, M. B. Schaefer, & D. Ness (Eds.), *Child parent research reimagined* (pp. 1-37). Brill. https://doi.org/10.1163/9789004421721_001
- Abrams, S. S., Schaefer, M. B., Ness, D., Abrams, C., Kurpis, M., & Ness, E. (2017, February 24-25). *Children as co-researchers of their digital activities* [Paper Presentation]. The 38th Annual Ethnography in Education Research Forum, University of Pennsylvania, Philadelphia, PA, United States.
- Al-Rodhan, N. R. F. (2009). *Sustainable history and the dignity of man: A philosophy of history and civilization triumph*. LIT.
- American Psychological Association. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, 51, 201-238. <https://doi.org/10.1037/h0053479>
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1966). *Standards for educational and psychological tests and manuals*. American Psychological Association.
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1974). *Standards for educational and psychological tests*. American Psychological Association.
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. American Psychological Association.
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. American Psychological Association.
- American Psychological Association, American Educational Research Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Psychological Association.
- Baez, B. (2007). Thinking critically about the 'critical' quantitative research as social critique. In F. K. Stage (Ed.), *Using quantitative data to answer critical questions* (pp. 17-23). Jossey-Bass.

- Balderas, A., Palomo, M., Beardo, J. M. D., & Rube, I. R. (2012, June 1). *Qualitative assessment of wiki-based learning processes*. <https://core.ac.uk/display/19492777>
- Beck, S. W., Bonner, S., Filkins, S., Landrigan, C., & Yoon, H. (2020, November 11). *Expanding formative assessment for equity and agency*. National Council of Teachers of English. <https://ncte.org/statement/expanding-formative-assessment/>
- Bennett, J., & Grant, N. S. (2016, May 23). Using an ecomap as a tool for qualitative data collection in organizations. *New Horizons in Adult Education & Human Resource Development*, 28(2), 1-13. <https://doi.org/10.1002/nha3.20134>
- Brewer, C., Knoepfel, R. C., & Clark Lindle, J. (2014). Consequential validity of accountability policy: Public understanding of assessments. *Educational Policy*, 1-35. <https://doi.org/10.1177/0895904813518099>
- Brink, S. C. (2011). A historical perspective of testing and assessment including the impact of summative and formative assessment on student achievement. *Dissertations*, 409. <https://aquila.usm.edu/dissertations/409>
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, 3, 296-322. <https://doi.org/10.1111/j.2044-8295.1910.tb00207.x>
- Burnett, C., & Merchant, G. (2014). Points of view: Reconceptualizing literacies through an exploration of adult and child interactions in a virtual world. *Journal of Research in Reading*, 37(1), 36-50.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105. <https://doi.org/10.1037/h0046016>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Cambridge Dictionary. (2021a). *Assessment*. Retrieved August 25, 2021, from <https://dictionary.cambridge.org/dictionary/english/assessment>
- Cambridge Dictionary. (2021b). *Imagine*. Retrieved November 20, 2021, from <https://dictionary.cambridge.org/dictionary/english/imagine>
- Collins English Dictionary. (2021). *Assessment*. In *Collins English Dictionary*. Retrieved August 25, 2021, from <https://www.collinsdictionary.com/dictionary/english/assessment>
- Collins, K. M. T., Onwuegbuzie, A. J., Johnson, R. B., & Frels, R. K. (2013). Practice note: Using debriefing interviews to promote authenticity and transparency in mixed research. *International Journal of Multiple Research Approaches*, 7(2), 271-284. <https://doi.org/10.5172/mra.2013.7.2.271>
- Cope, B., & Kalantzis, M. (2000). *Multiliteracies: Literacy learning and the design of social futures*. Routledge.
- Corrigan, J. A., & Onwuegbuzie, A. J. (2023). Towards sampling designs that are transparent, rigorous, ethical, and equitable (TREE): Using a tree metaphor as a sampling meta-framework in mixed methods research. In C. N. Poth (Ed.), *The Sage handbook of mixed methods research design* (pp. 130-142). Sage.
- Cowan, C. P. (1988). Working with men becoming fathers: The impact of a couples group intervention. In P. Bornstein & C. P. Cowan (Eds.), *Fatherhood today: Men's changing role in the family* (pp. 276-298). John Wiley.
- Cronbach, L. J. (1947). Test "reliability": Its meaning and determination. *Psychometrika*, 12, 1-16. <https://doi.org/10.1007/BF02289289>

- Cuellar, I., Martinez, C., Jimenez, R., & Gonzalez, R. (1983). Clinical psychiatric case presentation: Culturally responsive diagnostic formulation and treatment in an Hispanic female. *Hispanic Journal of Behavioral Sciences*, 5(1), 93-103. <https://doi.org/10.1177/07399863830051006>
- Daley, C. E., & Onwuegbuzie, A. J. (2011). Race and intelligence: It's not a black and white issue. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of Intelligence* (pp. 293-306). Cambridge University Press.
- Daley, C. E., & Onwuegbuzie, A. J. (2020). Race and intelligence: It's not a black and white issue. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 373-394). Cambridge University Press.
- Davis, T. M., Sellers, P. A., & Johnston, J. M. (1988, March 1). The factor structure and internal consistency of the Self-Concept and Motivation Inventory: What face would you wear? Preschool/kindergarten form. *Educational and Psychological Measurement*, 48(1), 237-246. <https://doi.org/10.1177/001316448804800129>
- Delgado, R., & Stefancic, J. (2012). *Critical race theory: An introduction* (2nd ed.). New York University Press.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. Praeger.
- Denzin, N. K., & Lincoln, Y. S. (2011). Introduction: The discipline and practice of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Sage handbook of qualitative research* (4th ed., pp. 1-25). Sage.
- Erickson, F. (2018). A history of qualitative inquiry in social and educational research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (5th ed., pp. 36-65). Sage.
- Farrah, G. A., Milchus, N. J., & Reitz, W. (1968). *The self-concept and motivation inventory: What face would you wear? SCAMIN manual of direction*. Person-O-Metrics.
- Feagin, J. R., & Cobas, J. A. (2008). Latinos/as and White racial frame: The procrustean bed of assimilation. *Sociological Inquiry*, 78, 39-53. <https://doi.org/10.1111/j.1475-682X.2008.00220.x>
- Fetters M. D., & Freshwater, D. (2015). The 1 + 1 = 3 integration challenge. *Journal of Mixed Methods Research*, 9, 115-117. <https://doi.org/10.1177/1558689815581222>
- Forzani, E., Corrigan, J., & Slomp, D. (2020). Reimagining literacy assessment through a new literacies lens. *Journal of Adolescent and Adult Literacy*, 64(3), 351-355. <https://doi.org/10.1002/jaal.1098>
- Forzani, E., Dobbs, C. L., Leider, C. M., *Malik, E., *Gragg, M., *Greszyck, C., & *Maher, C. (2024). Students as partners: Using an equity-oriented critical assessment practices (CAPS) approach in reading to empower students and inform instruction. *The Reading Teacher*, 77(5), 767-764. <https://doi.org/10.1002/trtr.2282>
- Forzani, E., Corrigan, J., Slomp, D., & Randall, J. (2024). *Prioritizing equitable social outcomes with and for diverse readers: A conceptual framework for justice-oriented reading assessment*. [Unpublished manuscript]. Boston University Wheelock College of Education & Human Development.
- Franklin, C., & Jordan, C. (1995, May). Qualitative assessment: A methodological review. *Families in society: The Journal of Contemporary Human Services*, 76(5), 281-295. <https://doi.org/10.1177/104438949507600502>
- Frels, R. K., & Onwuegbuzie, A. J. (2012). Interviewing the interpretive researcher: An impressionist tale. *The Qualitative Report*, 17(Art. 60), 1-27. <http://www.nova.edu/ssss/QR/QR17/frels.pdf>

- Frost, N. (2016). *Practising research: Why you're always part of the research process even when you think you're not*. Palgrave.
- Frost, N., & Holt, A. (2014). Mother, researcher, feminist, woman: Reflections on "maternal status" as a researcher identity. *Qualitative Research Journal, 14*(2), 90-102. <https://doi.org/10.1108/QRJ-06-2013-0038>
- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). *MAIN: Multilingual Assessment Instrument for Narratives*. Zentrum für Allgemeine Sprachwiss.
- Galton, F. (1904). Eugenics: its definition, scope, and aims. *The American Journal of Sociology, X*(1), 82. <https://doi.org/10.1038/070082a0>
- Gee, J. P. (1989). Literacy, discourse, and linguistics: Introduction. *Journal of Education, 171*(1), 5-17.
- Gee, J. P. (1996). *Social linguistics and literacies: Ideology in discourses* (2nd ed.). Taylor & Francis.
- Gee, J. P. (2000). Teenagers in new times: A new literacy studies perspective. *Journal of Adolescent & Adult Literacy, 43*(5), 412-420.
- Gee, J. P. (2011). *Social linguistics and literacies: Ideology in discourses* (4th ed.). Routledge.
- Gerber, H. R., Abrams, S. S., Curwood, J. S., & Magnifico, A. (2017). *Conducting qualitative research of learning in online spaces*. Sage.
- Gerber, H. R., Abrams, S., Onwuegbuzie, A. J., & Benge, C. L. (2014). From Mario to FIFA: What qualitative case study research suggests about games-based learning in a US classroom. *Educational Media International, 51*, 16-34. <https://doi.org/10.1080/09523987.2014.889402>
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*, 255-274. <https://doi.org/10.3102/01623737011003255>
- Hall, J., Gaved, M., & Sargent, J. (2021). Participatory research approaches in times of Covid-19: A narrative literature review. *International Journal of Qualitative Methods, 20*. <https://doi.org/10.1177/16094069211010087>. <https://journals.sagepub.com/doi/full/10.1177/16094069211010087>
- Hartman, A. (1978). Diagrammatic assessment of family relationships. *Social Casework, 59*, 465-476. <https://doi.org/10.1177/104438947805900803>
- Hartman, A. (1995). Diagrammatic assessment of family relationships. *Families in Society, 76*, 111-122. <https://doi.org/10.1177/104438949507600207>
- Henry, L. A. (2007). *Exploring new literacies pedagogy and online reading comprehension among middle school students and teachers: Issues of social equity or social exclusion?* (Doctoral dissertation). digitalcommons.uconn.edu/dissertations/AAI3282520
- Hertz, M. R. (1934). The reliability of the Rorschach ink-blot test. *Journal of Applied Psychology, 18*(3), 461-477. <https://doi.org/10.1037/h0071281>
- Hitchcock, J. H., & Onwuegbuzie, A. J. (Eds.). (2022). *Routledge handbook for advancing integrating in mixed methods research*. Routledge.
- Hitchcock, J. H., Onwuegbuzie, A. J., & Khoshaim, H. B. (2015). Examining the consequential validity of standardized examinations via public perceptions: A review of mixed methods survey design considerations. *International Journal of Multiple Research Approaches, 1-16*. <https://doi.org/10.1080/18340806.2015.1076757>
- Huot, B. (2002). (Re)Articulating writing assessment for teaching and learning. *All USU Press Publications, 137*. https://digitalcommons.usu.edu/usupress_pubs/137
- Johnson, R. B. (2012). Dialectical pluralism and mixed research. *American Behavioral Scientist, 56*(6), 751-754. <https://doi.org/10.1177/0002764212442494>

- Johnson, R. B. (2017). Dialectical pluralism: A meta-paradigm whose time has come. *Journal of Mixed Methods Research, 11*, 156-173. <https://doi.org/10.1177/1558689815607692>
- Johnson, R. B., Onwuegbuzie, A. J., Tucker, S., & Icenogle, M. L. (2014). Conducting mixed methods research: Using dialectical pluralism and social psychological strategies. In P. Leavy (Ed.), *Oxford handbook of qualitative research methods* (pp. 557-578). Oxford University Press.
- Johnson, R. B., & Gray, R. (2010). A history of philosophical and theoretical issues for mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Sage handbook of mixed methods in social and behavioral research* (2nd ed., pp. 69-94). Sage.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*, 112-133. <https://doi.org/10.1177/1558689806298224>
- Justice, L. M., Bowles, R., Pence, K., & Gosse, C. (2010). A scalable tool for assessing children's language abilities within a narrative context: The NAP (Narrative Assessment Protocol). *Early Childhood Research Quarterly, 25*(2), 218-234. <https://doi.org/10.1016/j.ecresq.2009.11.002>
- Kress, G., & Jewitt, C. (2003). Introduction. In C. Lankshear, M. Knobel, C. Bigum, & M. Peters (Series Eds.) & C. Jewitt & G. Kress (Vol. Eds.), *New literacies and digital epistemologies: Multimodal literacy* (Vol. 4, pp. 1-18). Peter Lang.
- Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2*, 151-160. <https://doi.org/10.1007/BF02288391>
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.
- Ladson-Billings, G. (2014). Culturally relevant pedagogy 2.0: Aka the remix. *Harvard Educational Review, 84*(1), 74-84. <https://doi.org/10.17763/haer.84.1.p2rj131485484751>
- Ladson-Billings, G. (2017). The (r) evolution will not be standardized: Teacher education, hip hop pedagogy, and culturally relevant pedagogy 2.0. In D. Paris & H. S. Alim (Eds.), *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.
- Lee, C. D. (2016). Influences of the experience of race as a lens for understanding variation in displays of competence in reading comprehension. In P. Afflerbach, (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (286-304). Taylor & Francis.
- Leech, N. L., & Onwuegbuzie, A. J. (2008). Debriefing. In L. M. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 199-201). Sage.
- Leu, D. J., Forzani, E., Rhoads, C., Maykel, C., Kennedy, C., & Timbrell, N. (2015). The new literacies of online research and comprehension: Rethinking the reading achievement gap. *Reading Research Quarterly, 50*(1), 37-59. <https://doi.org/10.1002/rrq.85>
- Logli, C. (2020). Culturally responsive assessment 2.0: Revisiting the quest for equity and quality in student learning. *Research & Practice in Assessment, 14*, 19-31. https://www.rpajournal.com/dev/wp-content/uploads/2021/09/A2_2.pdf
- Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal, 29*(6), 861-878. <https://doi.org/10.1080/0141192032000137349>

- Malinowski, B. (1967). *A diary in the strict sense of the term* (N. Guterman, Trans.). Harcourt, Brace & World.
- McReynolds, P., & Ludwig, K. (1987). 1987. On the history of rating scales. *Personality and Individual Differences, 8*, 281-283. [https://doi.org/10.1016/0191-8869\(87\)90188-7](https://doi.org/10.1016/0191-8869(87)90188-7)
- Merriam-Webster. (n.d.). Assessment. In *Merriam-Webster.com dictionary*. Retrieved August 23, 2021, from <https://www.merriam-webster.com/dictionary/assessment>
- Mertens, D. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 135-164). Sage.
- Mertens, D. M. (2007). Transformative paradigm: Mixed methods and social justice. *Journal of Mixed Methods Research, 1*, 212-225. <http://dx.doi.org/10.1177/1558689807302811>
- Mertens, D. M. (2010). Philosophy in mixed methods teaching: The transformative paradigm as illustration. *International Journal of Multiple Research Approaches, 4*, 9-18. <http://dx.doi.org/10.5172/mra.2010.4.1.009>
- Mertens, D. M., Bledsoe, K. L., Sullivan, M., & Wilson, A. (2010). Utilization of mixed methods for transformative purposes. In A. Tashakkori & C. Teddlie (Eds.), *Sage Handbook of mixed methods in social and behavioral research* (2nd ed., pp. 193-214). Sage.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23. <https://doi.org/10.3102/0013189X023002013>
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). Macmillan.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749. <http://dx.doi.org/10.1037/0003-066X.50.9.741>
- Meyer, A. E. (1965). *Educational history of the western world*. McGraw Hill.
- Mislevy, R. J. (2016). How development in psychology and technology challenge validity argumentation. *Journal of Educational Measurement, 53*(3), 265-292. <http://dx.doi.org/10.1111/jedm.12117>
- Mislevy, R. J. (2019). Advances in measurement and cognition. *The ANNALS of the American Academy of Political and Social Science, 683*, 164-182. <https://doi.org/10.1177/0002716219843816>
- Moll, L., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31*(2), 132-141. <https://doi.org/10.1080/00405849209543534>
- Montenegro, E., & Jankowski, N. A. (2017, January). *Equity and assessment: Moving towards culturally responsive assessment* (Occasional Paper No. 29). University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).
- Morrow, R. A., & Brown, D. D. (1994). *Critical theory and methodology* (Contemporary Social Theory). Sage.
- Mullings, B. (1999). Insider or outsider, both or neither: Some dilemmas of interviewing in a cross-cultural setting. *Geoforum, 30*, 337-350. [https://doi.org/10.1016/S0016-7185\(99\)00025-1](https://doi.org/10.1016/S0016-7185(99)00025-1)
- Natesan, P., Onwuegbuzie, A. J., Hitchcock, J., & Newman, I. (2019). Fully Integrated Bayesian thinking: A mixed methods approach to the 1 + 1 = 1 formula. *AERA Division D Newsletter, 10-12*. http://www.aera.net/Portals/38/docs/DivD/DNews_current/DivDNewsletter_Spring19.pdf

- Neimeyer, R. A. (1993). Constructivist approaches to the measurement of meaning. In G. J. Neimeyer (Ed.), *Constructivist assessment* (pp. 58-103). Sage.
- Newman, I., Onwuegbuzie, A. J., & Hitchcock, J. H. (2015). Using the general linear model to facilitate the full integration of qualitative and quantitative analysis: The potential to improve prediction and theory building and testing. *General Linear Model Journal*, *41*(1), 12-28. http://www.glmj.org/archives/articles/Newman_v41n1.pdf
- Onwuegbuzie, A. J. (2017, March). *Mixed methods is dead! Long live mixed methods!* Invited keynote address presented at the Mixed Methods International Research Association Caribbean Conference at Montego Bay, Jamaica.
- Onwuegbuzie, A. J. (2020). Child-parent research: Towards an ethical process for avoiding being PRICED out of research. In S. S. Abrams, M. B. Schaefer, & D. Ness (Eds.), *Child parent research reimaged* (pp. 205-234). Brill.
- Onwuegbuzie, A. J. (2021). Dear White Academic Colleague: Black Faculty Lives Matter: Towards avoiding the crime of silence. *International Journal of Multiple Research Approaches*, *13*(1), 98-109. <https://doi.org/10.29034/ijmra.v13n1postsc-ript>
- Onwuegbuzie, A. J. (2022, October). *Mixed methods research—where do we go from here?: The role of meta-methods research in the fourth and fifth industrial revolutions.* Inaugural Dr. Vimala Judith Kamalodeen lecture presented to students and faculty on behalf of the Mixed Methods International Research Association—Caribbean Chapter in collaboration with the Caribbean Journal of Mixed Methods Research.
- Onwuegbuzie, A. J., & Abrams, S. S. (2021). Nonverbal communication analysis as mixed analysis. In A. J. Onwuegbuzie & R. B. Johnson (Eds.), *The Routledge reviewer's guide to mixed methods analysis* (pp. 239-258). Routledge.
- Onwuegbuzie, A. J., & Abrams, S. S. (in press). *An integrated mixed methods approach to nonverbal communication data: A practical guide to collection and analysis in online and offline spaces.* Routledge.
- Onwuegbuzie, A. J., Abrams, S. S., & Forzani, E. (in press-a). The many SIDES of critical dialectical pluralism: A meta-philosophy—comprising a research philosophy, educational philosophy, and life philosophy—for addressing social justice, inclusion, diversity, and equity, and social responsibility. *International Journal of Multiple Research Approaches*.
- Onwuegbuzie, A. J., Abrams, S. S., Forzani, E., & Natesan Batley, P. (in press-b). A meta-framework to Re-IMAGINE the increase of Black methodologists in education research: Integrative mixed methods antiracist groundwork for investigating and nurturing equity. *International Journal of Multiple Research Approaches*, *14*(1).
- Onwuegbuzie, A. J., & Corrigan, J. A. (2021). Intra-study matching considerations when using mixed methods-based research approaches: A critical dialectical pluralistic approach. *International Journal of Multiple Research Approaches*, *13*(2), 116-136. <https://doi.org/10.29034/ijmra.v13n2editorial2>
- Onwuegbuzie, A. J., & Daley, C. E. (2001). Racial differences in IQ revisited: A synthesis of nearly a century of research. *Journal of Black Psychology*, *27*(2), 209-220. <https://doi.org/10.1177/0095798401027002004>

- Onwuegbuzie, A. J., Daniel, L. G., & Collins, K. M. T. (2009). A meta-validation model for assessing the score-validity of student teacher evaluations. *Quality & Quantity: International Journal of Methodology*, *43*, 197-209. <https://doi.org/10.1007/s11135-007-9112-4>
- Onwuegbuzie, A. J., Forzani, E., & Abrams, S. S. (2022). *History of quantitative, qualitative, mixed methods research, and educational assessment: A review*. Unpublished manuscript, University of Cambridge, Cambridge, England.
- Onwuegbuzie, A. J., Forzani, E., & Corrigan, J. A. (2023). Integrated mixed methods design-based research: A critical dialectical pluralistic approach. *Caribbean Journal of Mixed Methods Research*, *3*(1), 1-36.
- Onwuegbuzie, A. J., Forzani, E., Hitchcock, J. H., & Corrigan, J. A. (2022, August). *A meta-framework for conducting mixed methods impact evaluations: A critical dialectical pluralistic approach* [Paper presentation]. Online Mixed Methods International Research Association global conference.
- Onwuegbuzie, A. J., & Frels, R. K. (2013). Introduction: Towards a new research philosophy for addressing social justice issues: Critical dialectical pluralism 1.0. *International Journal of Multiple Research Approaches*, *7*, 9-26. <https://doi.org/10.5172/mra.2013.7.1.9>
- Onwuegbuzie, A. J., & Frels, R. K. (2015). A framework for conducting critical dialectical pluralist focus group discussions using mixed research techniques. *Journal of Educational Issues*, *1*, 159-177. <https://doi.org/10.5296/jei.v1i2.8662>. <http://www.macrothink.org/journal/index.php/jei/article/view/8662/7086>
- Onwuegbuzie, A. J., & Frels, R. K. (2016). *Seven steps to a comprehensive literature review: A multimodal and cultural approach*. Sage.
- Onwuegbuzie, A. J., & Hitchcock, J. H. (2019a). Toward a fully integrated approach to mixed methods research via the 1 + 1 = 1 integration approach: Mixed Research 2.0. *International Journal of Multiple Research Approaches*, *11*, 7-28. <https://doi.org/10.29034/ijmra.v11n1editorial1>
- Onwuegbuzie, A. J., & Hitchcock, J. H. (2019b). Using mathematical formulae as proof for integrating mixed methods research and multiple methods research approaches: A call for multi-mixed methods and meta-methods in a mixed research 2.0 era. *International Journal of Multiple Research Approaches*, *11*, 213-234. <https://doi.org/10.29034/ijmra.v11n3editorial2>
- Onwuegbuzie, A. J., & Hitchcock, J. H. (2022). Towards a comprehensive meta-framework for full integration in mixed methods research. In J. H. Hitchcock & A. J. Onwuegbuzie (Eds.), *Routledge handbook for advancing integration in mixed methods research* (pp. 565-606). Routledge.
- Onwuegbuzie, A. J., Hitchcock, J. H., Natesan, P., & Newman, I. (2018). Using fully integrated Bayesian thinking to address the 1 + 1 = 1 integration challenge. *International Journal of Multiple Research Approaches*, *10*(1), 666-678. <https://doi.org/10.29034/ijmra.v10n1a43>
- Onwuegbuzie, A. J., & Johnson, R. B. (Eds.). (2021). *The Routledge reviewer's guide to mixed methods analysis*. Routledge.
- Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. T. (2008). Interviewing the interpretive researcher: A method for addressing the crises of representation, legitimation, and praxis. *International Journal of Qualitative Methods*, *7*(4), 1-17. <https://doi.org/10.1177/160940690800700401>
- Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. T. (2010). Innovative data collection strategies in qualitative research. *The Qualitative Report*, *15*(3), 696-726. <http://www.nova.edu/ssss/QR/QR15-3/onwuegbuzie.pdf>

- Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. T. (2011). Innovative qualitative data collection techniques for conducting literature reviews. In M. Williams & W. P. Vogt (Eds.), *The Sage handbook of innovation in social research methods* (pp. 182-204). Sage.
- Onwuegbuzie, A. J., Rosli, R., Ingram, J. M., & Frels, R. K. (2014). A critical dialectical pluralistic examination of the lived experience of women doctoral students. *The Qualitative Report*, 19(5), 1-35. <http://www.nova.edu/ssss/QR/QR19/onwuegbuzie5.pdf>
- Paris, D., & Alim, H. S. (2014). What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harvard Educational Review*, 84(1), 85-100. <https://doi.org/10.17763/haer.84.1.982l873k2ht16m77>
- Paris, D., & Alim, H. S. (Eds.). (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.
- Patel, L. (2016). *Decolonizing educational research: From ownership to answerability*. Routledge.
- Pearson, P. D., Valencia, S. W., & Wixson, K. (2014). Complicating the world of reading assessment: Toward better assessments for better teaching. *Theory Into Practice*, 53, 236-246. <https://doi.org/10.1080/00405841.2014.916958>
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. https://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1465.pdf
- Randall, J., Poe, M., Slomp, D., & Oliveri, M. E. (2022). Disrupting white supremacy in assessment: Toward a justice-oriented antiracist validity framework. *Educational Assessment*. Advance online publication. <https://doi.org/10.1080/10627197.2022.2042682>
- Rosaldo, R. (1989). *Culture and truth: The remaking of social analysis*. Beacon.
- Saini, A. (2019). *Superior: The return of race science*. Beacon.
- Schaefer, M. B., Abrams, S. S., Ness, D., Kurpis, M., Abrams, C., & Ness, E. (2018, October 25-27). *What middle grades students discovered while researching with parents* [Roundtable presentation]. Annual Conference for Middle Level Education, Orlando, FL, United States.
- Schaefer, M. B., Abrams, S. S., Kurpis, M., Abrams, M., & Abrams, C. (2020). "Making the unusual usual": Students' perspectives and experiences of learning at home during the COVID-19 pandemic. *Middle Grades Review*, 6(2). <https://scholarworks.uvm.edu/mgreview/vol6/iss2/8>
- Schaefer, M. B., Abrams, S. S., Kurpis, M., Abrams, C., & Abrams, M. (2021). Pandemic meaning making: Messing toward motet. *English Teaching: Practice & Critique*, 20(2), 261-274. <https://doi.org/10.1108/ETPC-07-2020-0073>
- Schneider, P., Dubé, R., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*; University of Alberta Faculty of Rehabilitation. <http://www.rehabresearch.ualberta.ca/enni/about>
- Schön, D. (1983). *The reflective practitioner*. Basic Books.
- Spradley, J. P. (1979). *The ethnographic interview*. International Thomson Publishing.
- Spradley, J. P. (1980). *Participant observation*. Holt, Ainehart, and Winston.
- Starzewski, H., & Lloyd, H. (1999). Using the Stories/Narrative Assessment Procedure (SNAP) to monitor language and communication changes after a cochlear implant: A case study, *Deafness & Education International*, 1(3), 137-154, <https://doi.org/10.1179/146431599790561325>
- Stage, F. K. (2007). Moving from probabilities

- to possibilities: Tasks for quantitative criticalists. In F. K. Stage (Ed.), *Using quantitative data to answer critical questions* (pp. 95-100). Jossey-Bass.
- Strayer, G. D. (1913). *Standards and tests for measuring the efficiency of schools or systems of schools*. Bulletin, United States Bureau of Education. Whole No. 13: Report of the Committee of the National Council of Education.
- Strong, C. J., Mayer, M., & Mayer, M. (1998). *The Strong Narrative Assessment Procedure (SNAP)*. Thinking Publications.
- Taylor, E. (2006). A critical race analysis of the achievement gap in the United States: Politics, reality, and hope. *Leadership and Policy in Schools, 5*, 71-87. <https://doi.org/0.1080/15700760600499025>
- Teddlie, C., & Johnson, R. B. (2009). Methodological thought since the 20th century. In C. Teddlie & A. Tashakkori, *Foundations of mixed methods research: Integrating quantitative and qualitative techniques in the social and behavioral sciences* (pp. 62-82). Sage.
- Teranishi, R. T. (2007). Race, ethnicity, and higher education policy: The use of critical quantitative research. In F. K. Stage (Ed.), *Using quantitative data to answer critical questions* (pp. 37-49). Jossey-Bass.
- Tracy, E. M., & Whittaker, J. K. (1990). The social network map: Assessing social support in clinical practice. *Families in Society, 71*, 461-470. <https://doi.org/10.1177/104438949007100802>
- Tucker, S., Johnson, R. B., Onwuegbuzie, A. J., & Icenogle, M. L. (2020). Conducting mixed methods research: Using dialectical pluralism and social psychological strategies. In P. Leavy (Ed.), *Oxford handbook of qualitative research methods* (2nd ed., pp. 836-875). Oxford University Press.
- Vidich, A. J., & Lyman, S. M. (1994). Qualitative methods: Their history in sociology and anthropology. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 23-59). Sage.
- Vidich, A. J., & Lyman, S. M. (2000). Qualitative methods: Their history in sociology and anthropology. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 37-84). Sage.
- Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures*. Rand McNally.
- Weinbaum, R. K., & Onwuegbuzie, A. J. (2016). Getting more out of your interview data: Toward a framework for debriefing the transcriber of interviews. *Journal of Educational Issues, 2*, 248-264. <https://doi.org/10.5296/jei.v2i1.9216>
- Williams, R. (2000). Social Darwinism. In J. Offer (Ed.). *Herbert Spencer: Critical assessment* (pp. 186-199). Routledge.
- Wilson, A. D., Onwuegbuzie, A. J., & Manning, L. P. (2016). Using paired depth interviews to collect qualitative data. *The Qualitative Report, 21*, 1549-1573. <http://nsuworks.nova.edu/tqr/vol21/iss9/1>
- Witte, S., Bass, B., O'Byrne, W. I., Price-Dennis, D., & Sibberson, F. (2019, November 7). *Definition of literacy in a digital age*. National Council of Teachers of English. <https://ncte.org/statement/nctes-definition-literacy-digital-age/>
- Yancey, K. B., McCraw, B., & Filkins, S. (2018, October 25). *Literacy assessment: Definitions, principles, and practices*. National Council of Teachers of English. <https://ncte.org/statement/assessmentframingst/>
-