

Development and Preliminary Analysis of a Rubric for Culturally Responsive Research

The Journal of Special Education
2014, Vol. 47(4) 203–216
© Hammill Institute on Disabilities 2012
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0022466912436397
journalofspecialeducation
.sagepub.com



Audrey A. Trainor¹ and Aydin Bal¹

Abstract

Researchers and practitioners have struggled to promote optimal academic, behavioral, and postschool outcomes for historically marginalized youth from culturally and linguistically diverse backgrounds. While there is a growing body of evidence-based interventions in special education, the extent to which these interventions are culturally responsive remains unexplored. Culturally responsive research (CRR) has gained increased attention in social sciences. The authors developed a 15-item rubric to evaluate the cultural responsiveness of research. They applied the rubric to six studies in transition education identified as high-quality intervention studies to determine the extent to which these met the criteria for CRR. Results from this analysis demonstrated that while none of the studies were indicative of CRR across all rubric items, strengths in question relevancy, sampling, participant description, and data collection strategies were noted.

Keywords

research methods, cultural and linguistic diversity, experimental design, transition

Experimental and quasiexperimental research offers the field direction in the identification of effective instruction (i.e., what works, evidence-based practices). Federal education policies such as the Individuals With Disabilities Act of 2004 (IDEA) and the No Child Left Behind Act of 2001 (NCLB) contain mandates for the use of research-based decision making and instruction. The basic argument is that we must identify results-oriented, generalizable practices that are effective and timely to address common educational needs across a range of settings to maximize student outcomes. Evidence-based education is part of an ongoing reform movement to increase the knowledge and skills of U.S. students with economy and efficiency (Schneider & Kessler, 2007). Slavin (2002) argued that despite the potential costliness of randomized experiments, they are the pre-eminent way to establish generalizable, causal evidence in support of educational interventions.

Yet the evidence on which effective special education practices are based is only as sound as the methods used to create it. In recognition of this, prominent special education scholars identified quality indicators for a comprehensive, although not exhaustive, range of research designs. Quality indicators for experimental, quasiexperimental, single-subject, correlational, and qualitative designs were published in 2005. Other research designs, when aligned with the research question being asked, are legitimate tools for the knowledge production (Odom et al., 2005), but many researchers and policy makers assert that randomized

controlled experiments (RCE) are the most rigorous. Therefore, the evidence that results from RCE implementation is considered to be the most valid, reliable, and thus, generalizable, when identifying effective interventions (Cook, Tankersley, & Landrum, 2009; Gersten et al., 2005; National Research Council [NRC], 2002, 2005; Odom et al., 2005; Slavin, 2002, 2003).

In 2009, prominent scholars again devoted a special issue of *Exceptional Children* to the examination of the knowledge base in special education research in the key areas of reading, math, writing, and behavior. The conceptualization of evidence-based practices in this issue aligned with the quality indicators for experimental and quasiexperimental research previously published in the aforementioned 2005 special issue. Collectively, these articles presented compelling arguments and substantive evidence to demonstrate the contributions of experiments and quasiexperimental designs. But a vexing question remains: Do evidence-based practices hold promise for youth with disabilities for whom we have yet to consistently and systematically help achieve optimal academic, behavioral, and positive postschool outcomes? Here, we frame our concern as one about diversity and youth who experience

¹University of Wisconsin–Madison, USA

Corresponding Author:

Audrey A. Trainor, University of Wisconsin, 1000 Bascom Mall, Room 438, Madison, WI 53706, USA.
E-mail: aatrainor@wisc.edu

bias, discrimination, and/or are identified as members of groups who have historically been marginalized.

Scholars addressing persistent problems of equity in special education have argued that the extant scholarship in special education has yet to amass robust knowledge that addresses the strengths and needs of the racial/ethnic, linguistic, and economic diversity present in U.S. classrooms and society (Artiles, Kozleski, Trent, Osher, & Ortiz, 2010; August & Hakuta, 1998; NRC, 2002; Ortiz & Yates, 2010). Others have repeatedly called attention to the importance of viewing research as a situated cultural practice, acknowledging that from question inception to dissemination, all theoretical approaches and stages of inquiry are “culturally and socially mediated and negotiated” (Arzubiaga, Artiles, King, & Harris-Murri, 2008, p. 310). In 2003, the American Psychological Association (APA) published guidelines for conducting culturally responsive practice and research in psychology, highlighting implications for research design and implementation, and later revised this work (APA, 2010) to include consideration for cultural and linguistic diversity in assessment and research participation.

Despite these calls and guidelines for increased attention in educational research to culture, the concept has not been succinctly defined, nor have its implications been clearly outlined, in the context of special education research. What is the meaning of culture and its implications to research design, implementation, and dissemination? Our aim herein is threefold: First, we frame an argument for the importance of this question in the context of special education group experimental research. We review the extant literature and what is known about the conceptualization of culture for the purpose of developing criteria with which to evaluate research. Next, we describe the method we used to create a rubric with which to answer this key question: “To what extent does this group experimental study both incorporate and extend culturally responsiveness in research implementation and implications for practice and further study?” Last, we apply the rubric to a set of group experimental studies in transition education determined by Test and colleagues (2009) to have generated strong evidentiary support. Finally, we discuss findings and implications for special education research and practice.

Conceptualizing Culture

There are multiple and competing definitions of culture among and within social sciences. In this attempt to operationalize culturally responsive research (CRR), we tethered our conceptualization of culture to the philosophical, theoretical, and empirical work in cultural psychology, anthropology, and educational sciences. We used a sociohistorical conceptualization of culture: culture is a historically unique configuration, the residue of collective problem-solving activities among members of a social group in its efforts to

survive, adapt, and thrive within its everchanging social, economic, and physical contexts (Bal, 2011; Gallego, Cole, & the LCHC, 2001). This social inheritance is embedded in ideal and material artifacts and includes cognitive models (e.g., the scientific method as a framework for answering a research question) and material artifacts (e.g., the text-based professional journals as outlets for communicating research findings; Cole, 1996).

Culture is characteristically dynamic, multifaceted, and conflict laden, resulting in power/privilege differentiations, resistance, and innovations that are locally or heuristically accomplished (Bal, 2011). Cultural groups are heterogeneous, and individuals’ insider/outsider positions are negotiated and shifting in local contexts. Conducting CRR requires scholars to “strive to understand how people *assume*, but are also *given*, and *co-construct* multiple positions across contexts, depending on a host of forces that include local communities’ practices and history, as well as a person’s biographical trajectory” (Arzubiaga et al., 2008, p. 319).

Hence, our conceptualization of culture demands researchers to go beyond culture-blind and cultural deterministic approaches and to use a robust instrumental theory and methodology that considers complex interactions of individual, institutional, and interpersonal factors in the given context of conducting research. CRR moves beyond foci on a priori, static group identities (e.g., race or socioeconomic status) and outcomes (e.g., high school graduation) toward the inclusion of sociohistorical processes (e.g., special education referral and access to general education curricula) that locally reproduce enduring educational disparities.

The practices and tools associated with all research are imbued with beliefs, values, tools, and processes that have evolved historically and are situated in local, political, and sociohistorical contexts (Scheurich, 1997). The significance and implications of research results can be fully understood only if/when the physical, sociocultural, and historical contexts of the researchers and the participants frame the work. This is an important departure from the status quo because it acknowledges power and inequity as central players in the reproduction of educational disparities and, thus, variance in educational outcomes (Artiles et al., 2010).

Operationalizing CRR

While a strong rationale for producing CRR has been posited by key scholars concerned with the enduring problems of disproportionality and inequity in service delivery (see, for example, Artiles et al., 2010), operationalizing CRR is less developed. To identify potential criteria for evaluating CRR, we must draw from scholarship in psychology, education, and other social sciences. We now present findings from our literature review and identify relevant dimensions

of CRR. We categorize areas in need of attention and criterion development in three phases of inquiry: inception and design, implementation, and dissemination.

Inception and design. Contextual factors and cultural artifacts (e.g., models of human development and learning) shape researchers' thinking and planning as they develop investigations. Physical, cultural, and sociohistorical factors (e.g., location, demographics, experiences, beliefs, languages), at both group and individual levels, play some role in the way we conceptualize problems and design research questions (Ashing-Giwa, 2005). Theorizing inquiry requires researchers to formulate questions and hypotheses and to design studies, inseparable from their beliefs and values about inquiry, dominant practices in their academic fields, and available funding (Arzubiaga et al., 2008). For example, if RCE is considered a gold standard in research (Odom et al., 2005), scholars may begin by formulating experimental-type questions rather than first examining the complexity of the problem from multiple perspectives that include surveying gaps in the extant literature, thus violating the tenet that inquiry and its design be question driven (Phillips, 2006). Theorizing content is also embedded with valued practices from the researchers' fields, so constructs under examination may reflect a narrow, researcher-centric operationalization of a process that lacks ecological validity (Bernal, Bonilla, & Bellido, 1995). To avoid these potential pitfalls at the stage of conceptualization of method of inquiry and theoretical construct, CRR demands rigorous attention to the personal experiences, cultural practices, and assumptions researchers bring to a project in concert with the basic scientific principal of linking research to relevant theory (Arzubiaga et al., 2008).

Contextual factors influence relationships (i.e., access and opportunities) among and between researchers and participants (Fine, Weis, Weseen, & Wong, 2003; Merriam et al., 2001; Subedi, 2006). These factors sometimes create distance and other barriers, potentially encumbering research. For example, Ashing-Giwa (2005) cited the maltreatment of African American men in the Tuskegee Syphilis Study, a political and sociohistorical factor that damaged not only the lives of individual participants but also the relationship and trust between communities of researchers and participants in difficult-to-measure ways. Legislation regulating research may address humane treatment and begin to restore trust; however, the distal effects of prior discriminatory practices are difficult to measure. Members of some groups may be more likely than others to mistrust research activities and may therefore be less likely to participate. Similarly, researchers from the dominant culture may anticipate such mistrust and avoid attempts to include participant communities from historically marginalized backgrounds in research.

From inception and design, CRR attempts to minimize and address tensions associated with contextual factors by expanding the purpose of research and, therefore, the

research questions themselves, beyond the concerns of the researchers, their professional communities, and the private and governmental agencies they often serve. In other words, investigators using CRR designs promote the interests of those members of the investigatory community *and* those members of the participant community (Ladson-Billings & Tate, 2006). To attain this goal, CRR ideally includes members of participant community at all stages of research (Arzubiaga et al., 2008; Tillman, 2002). When this is not possible, investigators engaging in CRR must review the extant knowledge base and gain emic perspectives, taking care not to ignore critique and to acknowledge existing gaps (Wells, Merritt, & Briggs, 2009).

Implementation. In the initial stage of research implementation, issues of sampling arise. The implementation of RCE relies on randomization and the establishment of experimental and control groups. Sampling diverse populations in sufficient numbers for generalization continues to be a serious deficiency in the implementation of culturally responsive RCE and other designs (Calamaro, 2008). When participants from diverse backgrounds are omitted or included in insufficient numbers, narrow findings are interpreted too broadly (Graham, 1992). Also, replication, one affordance of RCE, has not been widely implemented to test hypotheses with specific subgroups of the population (Wells et al., 2009). Initial efforts to build trust and to include members of diverse communities as members of research teams are two strategies for increasing participants from historically marginalized groups, and thus potentially augment the cultural responsiveness of the work (Ashing-Giwa, 2005).

Building participants' trust, interest, and responsiveness through researchers' consideration of the strengths and needs of a diverse population requires transparency not only in sampling but also in intervention design and data collection (Wells et al., 2009). Of particular concern is the design of instruments and other intervention components that are "reliable, valid, and culturally consonant" (Ashing-Giwa, 2005, p. 134). This entails paying careful attention to intervention components' relevancy, measurement techniques, and language (Bernal et al., 1995). Intervention research, often implemented as RCE, is implicitly and explicitly connected to the researchers' aim to understand outcomes. In CRR, researchers consider the community-level outcomes, in addition to individual level outcomes, expanding the unit of analysis of inquiry (Chouinard & Cousins, 2007). These considerations may be conceptualized as contextual, that is, they include the analysis of the locales in which the intervention is likely to be implemented and the larger societal factors (Pope-Davis, Liu, Toporek, & Brittan-Powell, 2001).

Dissemination. The dissemination of research should result in its use (National Center for the Dissemination of Disability

Research, 1999). Practical uses of research may vary according to local contexts. For researchers, one typical use is to augment or sharpen the precision and reach of policy and to support organizational change. Attention to use in CRR requires that researchers also understand the local, political, and sociohistorical contexts in which people live, so that research implications for the broader society address local issues (APA, 2003). For example, as U.S. schools are becoming more racially/ethnically diverse, dissemination of evidence-based practices must include consideration of this issue. While research must still avoid extending implications beyond results, the dissemination of implications for practice and further research should be contextually and instrumentally framed, and any limitations in generalizability need to be acknowledged. In addition to publishing results in scholarly journals, presenting results to participants' communities using accessible language and focusing on the practical implications are essential (Ashing-Giwa, 2005). Next, we detail our synthesis of this comprehensive literature review and the resulting CRR rubric.

Developing a Rubric for CRR

Our work expands earlier efforts to identify quality indicators of research. Although both the earlier quality indicators and their application in rubric form do have relevance in the consideration of CRR, neither their original establishment in 2005 nor the application of these criteria to existing bodies of research in 2009 explicitly examined what aspects of culture might be important to consider in the evaluation of intervention research. Yet areas of overlap between our rubric and the previous aforementioned work exist. For example, a relevant, established rubric item applied to the examination of experimental reading research is, "Samples are comparable across conditions on relevant characteristics" (Chard, Ketterlin-Geller, Baker, Doabler, & Apichatabutra, 2009, p. 271). We concur and include a version of this indicator of CRR in our rubric. But we attempt to expand the issue of recruitment, making a concerted effort to consider additional variables that have implications for generating evidence about what works in addressing diverse preferences, experiences, strengths, and needs of youth with disabilities.

Our aim is to strengthen—not to replace—commonly accepted standards and quality indicators of experimental research. To do this, we followed the framework for quality indicators of the experimental studies outlined in the aforementioned scholarly papers. Basing our argument and rubric items on the principle that research is itself a cultural practice (Arzubiaga et al., 2008), our ultimate goal is to provide a conceptual tool to enhance researchers' reflexivity during conceptualization, design, implementation, and dissemination of research.

Review of the literature. We reviewed psychology, educational, and other social science literature to identify guidelines, rubrics, theoretical papers, research articles, and syntheses on CRR and related interventions. We searched three academic databases (ERIC, PsycINFO, and Google Scholar) for the following combinations of keywords: *culturally responsive*, *culturally competent*, *culturally adequate*, *cultural competency*, *cultural adequacy*, or *cultural responsiveness* and *research*. The searches were filtered for dates ranging from 2000 to 2010. We also manually searched the publications that we identified through the literature search and contacted experts on equity and diversity in education and psychology, seeking existing checklists or rubrics and related sources. We found no published rubric for evaluating CRR.

Next, we synthesized the resulting relevant theoretical papers, guidelines, and empirical studies that detailed the related tenets of cultural responsiveness in teaching, assessment, therapy, and empirical studies that highlighted strategies for conducting CRR. In addition to highlighting the significance of our rationale, the literature search formed the foundation of our rubric for CRR, providing a comprehensive interdisciplinary framework of criteria.

Rubric development. We organized findings from the literature review into domains for rubric item development following the standards for reporting on empirical research (American Educational Research Association [AERA], 2006). These domains are (a) problem formulation, (b) design and logic, (c) sources of evidence, (d) measurement/assessment process, (e) analysis and interpretation, and (f) dissemination. Although these domains represent dominant cultural beliefs and practices about research, this organization is appropriate for this preliminary examination of CRR because it is well understood by educational researchers. We created rubric items by reviewing the literature examining tenets associated with CRR and identifying criteria from the theoretical and empirical articles, book chapters, and guidelines. Following Chard et al.'s (2009) recommendation, we created ratings for the levels of rigor or fitness for each criterion. We used a 3-point Likert-type scale (0–2) to represent the variance in CRR ratings of criteria. Ratings of 0 indicate an absence of documentation of the role of culture, using a culture-blind approach. Ratings of 1 indicate documentation that culture was viewed as a categorical, static variable (e.g., race, social class), a determinant of participants' perceptions and behaviors. Ratings of 2 indicate documentation of a practice-oriented focus on the intersection of cultural and contextual (i.e., individual, institutional, and interactional) factors.

The final 15-item rubric is included in Table 1. Space limitations preclude a detailed discussion of each item; sources cited for each item in the rubric are available on request. Here, we illustrate the relationship between extant

Table 1. Rubric for Culturally Responsive Research.

Rubric item	Score		
	0	1	2
1. Foundational construct(s) of the study	The construct under examination is implied but not explicitly discussed.	The construct under examination is explicit but taken as universal based on a norm-referenced sample with dominant cultural/linguistic background. Evidence of alternative conceptualizations is not presented.	The construct under examination is addressed comprehensively and adequately; multiple perspectives and/or competing ideas are presented with a presentation of evidence of alternative conceptualizations.
2. Relevancy of the research problem	The relevancy of the research problem(s) to participants' interests and needs and context is not discussed.	The relevancy of the research problem is discussed, as it relates to the field and/or the researcher's interest or line of inquiry.	The relevancy of the research problem addresses both the researcher's line of inquiry and the participants' and local communities' interests and needs.
3. Critical and comprehensive review of the relevant literature	The review of extant literature results in a narrow rationale for the study that does not address what is known about the problem.	The review of extant literature includes scholarship as it relates to the research problem relevant to the unit of analysis.	The review of extant literature is critical and creates a dialogue with studies using alternative methodologies and perspectives on the research problem relevant to the unit of analysis.
4. Justification of the theoretical framework	The theoretical framework of the study is not discussed explicitly.	The theoretical framework of the study is discussed abstractly, only as it relates to the construct under examination.	The theoretical framework is discussed comprehensively as it relates to physical, sociocultural, and historical contexts of the study and participants' lives. The framework is justified with a critical examination of its limitations to study the problem, participants, and the contexts of the study.
5. Description of participants	Description of participants' demographic characteristics includes two or fewer characteristics (race, gender, income, disability).	Description of participants includes more than two characteristics; however, the description is limited to the dimension of the individual.	Description of participants includes both individual characteristics and the institutional dimensions (e.g., status, institutionalized social practices) for both the control and intervention groups.
6. Description of researchers and interventionists	Description of researchers and interventionists (e.g., teacher, translator) includes two or fewer individual characteristics (e.g., race, experience, and language).	Description includes more than two individual characteristics (e.g., race, economic background, gender, disability); however, the description is limited to the dimension of the individual.	Description includes individual characteristics and the contextualized institutional dimensions and relational positions among the participants and interventionists (e.g., power, status, and insider/outsider positions).
7. Description of sampling procedures	Recruitment and sampling methods are not discussed.	Recruitment and sampling methods are discussed but lack detail about the rationale for the exclusionary criteria (e.g., English language learners) and the congruence of participants' experiences and/or preferences (e.g., language preference).	Recruitment and sampling methods include differentiation based on participants' experiences and preferences, maximizing the potential to include diverse populations.

(continued)

Table 1. (continued)

Rubric item	Score		
	0	1	2
8. Description of research setting	Description of the research setting is not discussed.	Description of the research setting includes the school and community characteristics relevant to the construct under examination.	Description of the setting includes physical, sociocultural, and historical contextual factors and their interaction with the research process.
9. Description of data collection strategies	A rationale for the data collection strategies is not discussed.	A rationale for the data collection strategies is provided; however, it is limited to a technical discussion of the methodology.	A rationale for the data collection strategies includes consideration of participants' cultural and linguistic preferences, needs, and strengths. Multiple data collection methods are used to maximize accessibility (e.g., using instruments in multiple language, using participant-selected locales). Description includes discussion of interactions between the researchers and participants.
10. Ecology of the intervention	The intervention includes a contrived context, task, and control for variables to the extent that its application in real life is unlikely.	The intervention includes a context, task, and variables that generally represent participants' real life experiences yet the intervention aligns more closely with the research design.	The intervention is aligned with participants' experiences and/or preferences. The integrity of the participants' experiences and contexts is balanced with the researchers' design.
11. Intervention design	<i>Culture-free approach:</i> The intervention does not consider fundamental aspects of cultural and linguistic diversity that participants, interventionists, and researchers bring to the study. The diversity of the physical, sociocultural, and historical contexts is not discussed.	<i>Culturally sensitive approach:</i> The intervention incorporates procedures to incorporate individual and within-group diversity including a combination of the following: Training of interventionists for research with diverse groups; embracing participants' perspectives and practices, inviting community representation; ensuring availability and accuracy of translation and interpretation of intervention materials and procedures, considering the implications of legal issues; and/or examining the applicability of interventions to participants' lives.	<i>Culturally responsive approach:</i> The intervention study incorporates methods and procedures to address diversity but also meets all three fundamental criteria for culturally responsive educational interventions: to improve academic achievement, skills and knowledge, and social outcomes; to affirm participants' cultural and personal identities; and to facilitate the development of participants' critical perspectives both to develop an awareness of and capacity to challenge inequities that they experience.
12. Assessment of intervention efficacy	The validity, reliability, and language of the measurement tool(s) are not discussed.	The validity, reliability, and language of the measurement tool(s) are discussed, but the measurements are standardized and norm-referenced for a population other than the sample.	The validity, reliability, and language of the tools are inclusive of the populations representative of participants OR the limitation/lack of availability of such tools for the sample are discussed.

(continued)

Table 1. (continued)

Rubric item	Score		
	0	1	2
13. Presentation of findings	The results are not disaggregated according to the participant and setting characteristics.	The results are disaggregated according to participant characteristics between and within the intervention and control groups, but are limited to disability, race, income, or language.	The results are disaggregated according to participant characteristics between and within the intervention and control groups and include intersections of participant characteristics.
14. Analysis and interpretation	<i>Culture-blind approach:</i> Participants' cultural, linguistic, and economic backgrounds and contextual factors are not included in data analysis and interpretation.	<i>Cultural deterministic approach:</i> Participants' backgrounds and contextual factors are analyzed as categorical and static variables. Differences among the participants are interpreted based on the disadvantages associated with living conditions, demographic characteristics, or participants' lack of competencies in mainstream skills and knowledge.	<i>Cultural instrumentalist approach:</i> Participants' backgrounds, contextual, and cultural factors are analyzed as dynamic, complex, and dialogical. Differences within the participants are interpreted as situated in affordances and constraints of the physical, sociocultural, and historical relations of the context. Factors under consideration include organizational structures, power distribution, and participants' identities.
15. Discussion of dissemination	Dissemination strategies are limited to the presentation of data in the article.	Dissemination strategies extending beyond the article are discussed (e.g., the data were shared with teachers and families).	Dissemination strategies are strategically selected to maximize sharing of knowledge with clear practical benefits to participants' communities writ large.

literature, two select criteria, and the rating schema. Item 12, *Assessment of intervention efficacy*, addresses the somewhat straightforward issue of the validity and reliability of assessment tools. The significance of this criterion has long been acknowledged in educational research and practice: Assessment tools are artifacts whose participatory constructs, concepts, and language are inherently cultural (Solano-Flores, 2008). Despite tacit agreement among researchers and practitioners that assessments must be congruent with students' characteristics and experiences (e.g., language), appropriate use of culturally responsive assessment tools continues to be problematic in the identification of disability, preferences, strengths, and needs for youth from historically marginalized groups (Figueroa & Newsome, 2006). Thus, the selection of appropriate assessment measurements continues to be an explicitly stated tenet and criterion for CRR (AERA, 2006; APA, 2003, APA 2005). An embedded question in Item 12 is "Were reliability and validity of the assessments, particularly in regard to the norm-referenced sample upon which these are based, explicitly and intentionally matched to the participants' characteristics and experiences?" The ratings indicate the extent to which the detailed

description of assessments and measurements explicated validity, reliability, and cultural congruence with participants. A score of 0 indicates an absence of the documentation of validity, reliability, and the language of the assessment. A score of 1 indicates that, although these were discussed, the population on which the validity and reliability were normed did not represent the study's sample. When documentation of validity, reliability, and language of the measurements included the range of populations in the study's sample or when the lack of such instrumentation constituted an acknowledged limitation, a score of 2 was assigned.

Item 14, *Analysis and interpretation*, focuses on the conceptualization of culture during analysis of results. We anchor this criterion to arguments for the expansion and importance of culture as pertains to teaching and learning (Artiles et al., 2010) and professional guidelines for research (AERA, 2006; APA, 2003). Therefore, an embedded question is, "Was the role of culture conceptualized contextually and instrumentally accounting for the relevant individual, institutional, and interpersonal factors?" A score of 0 on this item indicates that analysis and interpretation did not include consideration of cultural and contextual factors, while a score of 1 indicates that

these variables are considered as categorical and static, without examining intersections of variables. A score of 2 indicates that the analysis of culture reflects the complexity of variables that may extend beyond categorical description, may vary within categories, and may change across contexts.

Once drafted, we submitted the rubric for review from two scholars, one a tenured professor with expertise in testing development and educational measurement, at our university's center for testing and evaluation, who were asked to provide critical feedback on the design of the set of items, as well as the rating structure for each item. Based on this feedback, we refined and delineated the purpose of the rubric; that is, we proceeded developing the CRR rubric as a tool to provide an evaluation of each criterion rather than to generate an overall score. This decision allowed flexibility, an important consideration, as reviewers' preliminary feedback demonstrated that the ratings for each criterion were not parallel or comparable across the rubric. Next, we submitted the rubric for a review of its content from four prominent special education scholars from other major research universities, each with expertise in research design and intervention with culturally and linguistically diverse populations. Three scholars accepted our request and provided feedback via telephone conference and an overall review of the rubric's content. Based on their feedback, we revised the rubric in two ways. According to one intervention scholar, we calibrated the scoring criterion of 2, the highest score, to include the acknowledgment of limitations of existing culturally responsive measurement instruments or limited participant sampling. In this way, we expanded the attainability of a score of 2 to include research that explicitly acknowledges such limitations as this demonstrates the reflexivity aligned with the theoretical tenets of cultural responsiveness. Following content revisions, we applied the rubric to an intervention study outside the targeted set of transition studies identified by Test and colleagues (2009) to calibrate the raters' interpretation of the scale. In this application, we practiced using the rubric, compared results, and noted further need for revisions, editing for parallel language and clarity, and made several adjustments to clarify the expectations of the levels of rigor associated with each score.

Rubric application. We (both authors) applied the rubric to a set of group experimental studies identified by Test and colleagues (2009) as rigorous and exemplary of evidenced-based practices in transition, our area of interest. We read each of the six studies and evaluated the work according to each rubric item. Following independently scoring the studies in the set, we compared scores and determined interrater reliability. Following Chard et al. (2009), the interrater reliability was calculated by dividing the number of exact matches on ratings at the component level by the total number of exact matches and disagreements. This resulted in a

reliability score of 0.6. Next, we discussed instances where there were differences, returned to the original texts for clarification, and negotiated evaluative judgment until 100% agreement was reached (Bazeley, 2007; Richards, 2005). For example, on Item 6, we reached initial 100% agreement that none of the articles presented a description of the researchers or interventionists, resulting in a score of 0 on this item for each article. On Item 5, however, our separate evaluations of the inclusion of participant characteristics were in agreement on five of six articles. Returning to the table and related text in the Sinclair, Christenson, and Thurlow (2005) study of school completion allowed us to verify and agree upon the score of 1, thus reaching interrater agreement.

Results

We applied our CRR rubric to a set of intervention studies in special education to examine the extent to which the body of work identified as evidence-based through the implementation of RCE also aligned with CRR. We chose a set of intervention studies that had previously been identified as rigorous transition research (Test et al., 2009); however, the rubric could be applied to any empirical study. To do this systematically and congruently with established quality evaluations of special education research, we created a rubric. To our knowledge, this is the first rubric that addresses CRR in special education. Table 2 provides scores for each rubric item by study. Analysis across rubric items illustrates constellations of scores (0s, 1s, and 2s) in each research domain, indicating a range of alignment with CRR. We acknowledge that the consideration of CRR has not previously been concretized as quality indicators for special education research in the detail presented herein. Thus, we expected to find both domains of research and individual items that did not meet these criteria simply because some indicators, representative of the tenets of CRR, have not been explicitly discussed or identified as quality indicators in our field. We see this analysis as an attempt to operationalize CRR research and one that will forward an important and perhaps transformative discussion.

Identifying Opportunities for CRR as Value Added

We begin our analysis with the most obvious pattern: None of the articles received the optimal score of 2, an indication that the tenets of CRR had been met, on any rubric item. Scores of 2 generally indicated documentation that culture was understood not as a stand-alone variable but as a dynamic and instrumental process. To illustrate, a score of 2 for Item 9, *Description of data collection strategies*, indicates the use of data collection methods that are responsive to participants' preferences and experiences (e.g., language

Table 2. Rubric Scores for Six Experimental Studies in Transition Research.

Rubric item	Bates et al. (2001)	Izzo et al. (2000)	Martin et al. (2006)	Nelson et al. (1994)	Sinclair et al. (2005)	Van Reusen & Bos (1994)	Total (%)		
							0	1	2
1. Construct	1	0	0	1	1	0	50	50	0
2. Relevancy	0	1	1	1	1	1	17	83	0
3. Literature	0	1	1	0	0	1	50	50	0
4. Theoretical	0	0	0	0	0	0	100	0	0
5. Participants	0	1	1	1	1	1	17	83	0
6. Researchers	0	0	0	0	0	0	100	0	0
7. Sampling	1	1	1	0	1	1	17	83	0
8. Setting	0	1	1	1	1	0	33	67	0
9. Data collect	1	1	1	0	1	1	17	83	0
10. Intervention	0	0	1	1	1	1	33	67	0
11. Intervention design	0	0	0	0	0	0	100	0	0
12. Assess intervention	0	0	0	1	0	0	83	17	0
13. Finding present	0	0	0	0	1	0	83	17	0
14. Analysis and Interpretation	0	0	0	0	0	0	100	0	0
15. Dissemination	0	0	0	0	0	0	100	0	0

preference) *and* that relevant details about the researcher/interventionist interactions with participants are documented. None of the articles provided this level of detail in the documentation of data collection.

For example, Martin and colleagues (2006) did not explain whether it was necessary to translate the survey instrument they used into Spanish despite that 3.7% of their participants identified as Hispanic/Latino. The authors do not discuss whether the survey items were designed to be congruent with participants' cultural conceptualizations of Individualized Education Program (IEP) participation or whether trusting, collegial relationships existed between teachers, parents, and youth whose interactions were the subject of study. At the same time, the extant literature documents the reluctance of some Latino and African American parents to assert opinions at IEP meetings, for a myriad of reasons that include feelings of being ignored or disrespected (Harry, 1992), deference to teacher expertise (Bailey, Skinner, Rodriguez, Gut, & Correa, 1999), and tacit disagreement (Kozleski et al., 2008). Because Martin and colleagues discussed the relevancy of their unit of analysis at the student level, provided a detailed description of participants and data collection strategies, and included contextual and controlled variables that generally represented participants' real life experiences, we agreed that their study showed promise for CRR, and we scored accordingly (see Table 2).

Notably, we did not find documentation warranting a score of 2 on any rubric item in this or any other study. Evidence of a score of 2 on this item for this study might

have included a discussion of the strategies the researchers used to create the survey instrument in Spanish and/or one that reliably operationalized cultural conceptualizations of IEP involvement, as well as strategies for developing trust with participants to increase the depth and breadth of their input. We interpret this finding as an indication that the tenets of CRR, as represented by the rubric, have not been explicitly addressed. One contributing factor may be the absence of operational definitions of CRR criteria on which there has been agreement and standardized practice. Our rubric provides a framework to address this underdeveloped area of special education research.

In fact, as we conducted our literature review, developed, and then applied our rubric, we identified dissonant themes when we simultaneously considered the tenets of CRR and those of RCE. RCE is a research methodology within the positivistic paradigm of unbiased approaches to inquiry; however, to conduct CRR requires an examination of potentially consequential biases and cultural practices. These ideas may create dissonance for researchers. What at first consideration may seem incongruent, however, may actually be an opportunity to expand established quality criteria. For example, Gersten and colleagues (2005) asserted that the professional qualifications and experiences align with the interventions they implement. We developed Item 6, *Description of researchers and interventionists*, by expanding this criterion so that it explicitly addresses culture. Item 6 requires the additional documentation of individual characteristics, relational positions, and contextualized dimensions of the interventionist. Because the goal of RCE is

often generalization, this expanded criterion lends more precision to generalizability. For example, Izzo, Cartledge, Miller, Growick, and Rutkowski (2000) intervened by extending transition services for 2 years beyond high school. While the remaining participants in the experimental group were more likely than those in the control group to successfully attain employment goals, 17 of the 79 participants in the experimental group left the study. On Item 6, we scored Izzo and colleagues' (2000) work at a 0 because they provide too little description of the interventionists (i.e., "job training coordinators," p. 148). Understanding more about the characteristics and contexts of the interventionists and their relationships to participants (i.e., criterion in Item 6) would contribute to our understanding of the ecological validity, and thus the efficacy, of the intervention. We understand that the attrition of participants could be linked to other factors, and we acknowledge that its effect cannot be fully understood without additional analyses by the researchers. Ultimately, though, this information is essential if we are to generalize findings.

Finding that all six studies omitted descriptions of the researchers/interventionists or the relational dimensions of research (100% scored 0 on Item 6) was not surprising. Controlling for researcher biases is a prevailing tenet of RCE designs. Although beyond the scope of this article, the theoretical link between nonbiased research and culture-blind research is important to note. Strategies for limiting potential biases and culture-blind approaches require some discussion of researcher/interventionists' perspectives and experiences that shape their approaches to research design and implementation. Rather than increasing biases, transparency of positionality contributes to a "culture-centered" approach prioritized in CRR (APA, 2003, p. 380). This latter position is connected to multicultural education principles that address equity by examining the lack of representation in the teaching force relative to youth and their families across school populations and the effect of this disproportionality on shared values, beliefs, and practices. Similarly, researchers and participants experience differences in power and status or differences in the benefits afforded by research, but these issues have not been fully considered in special education research. Our rubric is an evaluative tool that may lead to a deeper consideration of which researcher/interventionist factors and contexts are important to report to move toward CRR and to fulfill existing requirements of experimental research (Gersten et al., 2005).

All six articles received a score of 0 on 4 more of the 15 items (Items 4, 11, 14, and 15). Generally, a score of 0 indicated that there was insufficient documentation that the role of culture in these intervention studies was a key consideration, whether at the surface level of variable or at the more complicated level of a negotiated process. Again, some of these scores are indicative of a more traditional practice of research. For example, in Item 15, *Discussion of dissemination*, a 0 indicated

that dissemination beyond the article was not discussed. A tenet of CRR is that dissemination efforts reach stakeholders who are potentially affected by the results and implications (a score of 1 on the rubric). This requires sharing data and translating findings into practical knowledge for use in the local participant communities *and* the larger community of scholars and practitioners. Sinclair and colleagues' (2005) study of *Check & Connect*, a dropout prevention intervention, demonstrated that youth in the experimental group experienced increased attendance and decreased dropout rates. While the authors detail implications for practice and further research, they stop short of explaining how they shared this work with participants' communities or mobilized local resources to increase the reach of the intervention. We acknowledge that typical dissemination discussions infrequently include this information; however, increasing our expectations for dissemination has the potential to bolster connections we make within and across research settings. Proponents of CRR argue that the translation of research to practice and the sharing of findings with participants' communities not only increase the impact of program implementation but also may increase the recognition of people from historically marginalized communities of a vested interest in educational research (Ashing-Giwa, 2005; Chouinard & Cousins, 2007). While lack of documentation does not mean that such dissemination efforts were not completed, doing so provides an opportunity for accountability in research. Again, the rubric demonstrates the possibilities of CRR in special education contexts, operationalizing components not previously included in the repertoire for publication of traditional research and moving analysis toward CRR in incremental steps.

Improving CRR Through Strengthened Theoretical Frames and Sharpened Analyses

Some of the low-scoring rubric items coincided with standard requirements for research designs that are foundational to RCE. For example, Item 4, *Justification of the theoretical framework*, and Item 14, *Analysis and interpretation*, are common foci for researcher attention, irrespective of CRR. The difference or shift between familiar expectations and those in our CRR rubric is the focus on cultural complexities. For example, researchers in special education may commonly address the theoretical frameworks that undergird their studies, but a tenet of CRR is to do so explicitly attending to participants' and communities' local, political, and sociohistorical contexts. For example, Van Reusen and Bos (1994) examined the effects of an intervention designed to motivate students to participate in their IEP meetings, reporting that youth in the experimental group were able to identify goals and verbally contribute at rates higher than the youth in the control group. In their discussion of IEP meetings, however, they do not provide sufficient detail about the context of the IEP meetings nor

about the youth's backgrounds and experiences. Yet, existing communication theories indicate that context is an important factor of shared meaning through discourse and that culture is consistently a factor in interpersonal communication (Gudykunst & Lee, 2003). Therefore, we noted that an implicit, unidentified theoretical framework invoking the principles of collaborative decision making or communication was in operation, as was a similarly implicit cultural identity framework.

The tenets of CRR underscore the need for such theoretical frameworks to be made explicit and to aid in making the approach to research design and implementation more exacting, and potentially increasing the generalizability of the results. This is in alignment with the established criterion that researchers explicitly connect their work to existing theoretical frames (AERA, 2006; APA, 2003, 2005). Documentation warranting a score of 2 on this item on our rubric would include explaining the role culture might play in interactions when people (e.g., teachers, parents, students) have different preferences, experiences, strengths, and needs. While these researchers clearly document the legal imperative in the literature review, an explicit framework incorporating communication and interaction contexts and cultures would potentially strengthen the implementation of the study and its implications (i.e., replicability and practical implementation) by addressing how a mandate for student IEP involvement is made more complex in schools where parents, youth, and teachers may have disparate views about what it means to be involved, to assume responsibility, and to communicate needs and expectations. For example, relevant frameworks include theories of cultural psychology (Cole, 1996), cross-cultural interactions (Gudykunst & Lee, 2003), and capital theory (Lareau & Horvat, 1999).

Enhancing the CRR Tenets Within Existing Research Approaches

On many of the 15 items, we noticed that the target criteria of CRR, as articulated in the rubric, were within reach of the research presented in these special education studies. On one quarter of the items (Items 2, 5, 7, and 9), five of the six articles scored 1s, indicative of researchers' conceptualization of culture as a notable variable. In these cases, conceptualizations of culture were static and tightly coupled with participants', rather than researchers', characteristics and contexts, and did not reflect complex negotiated processes. For example, with the exception of Bates, Cuvo, Miner, and Korabel's (2001) study of community-based instructional interventions, the remaining five articles do report more than two characteristics associated with participants, most commonly disability, gender, and race/ethnicity, at the level of the individual.

We were unable to find, however, descriptions of complex individual or institutional dimensions (Item 5,

Description of participants) necessary for a score of 2 that augment the explanation of the results. Izzo and colleagues' (2000) documented participant characteristics across gender, race/ethnicity, and disability for both experimental and control groups, but they do not provide contextual descriptions that explain why participants' were identified as "at-risk for not maintaining employment" (p. 141). What were the institutionalized social practices that may have contributed to the problem? And, following the presentation of data (see Item 13), how might this description inform our understanding of the intervention's varied success? Expanding the description of participants' contexts has the potential to aid interpretation of results and implications. For example, was the intervention more likely to work with youth who lived in communities where adult employment was consistently maintained? Did young women of color and their male peers in the experimental group benefit similarly? We understand that such discrete analyses may have been omitted because participant numbers in each subgroup were insufficient for the statistical tests and measures. Still, thoroughly describing a study's sample is a common quality criterion in special education research (Chard et al., 2009; Montague & Dietz, 2009; Odom et al., 2005); thus, the rubric identifies an area in need of improvement that aligns not only with the tenets of CRR but also with the basic RCE concept of generalizability.

The first four rubric items address documentation of the research inception and planning. On 3 of the 4 items (Items 1, 2, and 3), scores of 1 were obtained by at least half of the six articles, indicating documentation approaching CRR. For example, the construct under examination (Item 1) as described by Nelson, Smith, and Dodd (1994) is made explicit, yet the pointed focus on handwritten job applications excludes a range of considerations for the employment application process, some of which are context embedded and require attention to culture as comprehensive and negotiated processes. Because CRR capitalizes on culture as an interactional process, a score of 2 would expand the conceptualization of applying for jobs noting that written applications provide but one point of entry into the job market. For example, an expanded discussion might include acknowledgment that in some contexts social networks requiring no written application are more common and/or that handwritten applications may be indicative of types of employment that are more or less relevant in the lives of the participants. While the expansion of the construct may seem to some to exceed the parameters of Nelson and colleagues' study, doing so, in accordance with CRR, augments the field's understanding of the construct under examination in relation to intended beneficiaries of the research. This is a salient issue because, although this study was published in 1994, a time when handwritten applications were common, this study was recently identified as an evidence-based transition intervention (Test et al., 2009) despite that other modes of applying for

positions (e.g., Internet-based application systems) might also be important to consider. Relevancy is a related issue (Item 2, *Relevancy of the research problem*). We scored Nelson et al.'s (1994) study as a 0 on the issue of relevancy because handwritten applications are becoming outmoded in the contemporary context. This second point is not a critique of the original work; rather, it is a reminder that context matters. As our scoring indicated, the construct under examination is narrow and its relevancy limited when viewed through the lens of CRR.

Implications

One common objective of school reform is to systematically address youth whose needs are not effectively addressed in school. Despite concerted effort by special education researchers to identify practices considered to be evidence-based, and thus effective, teachers still struggle to improve the achievement and postschool outcomes for youth from historically marginalized groups. The extent to which this persistent problem can be linked to the production of evidence through research is unclear. This examination of a sample set of studies identified as strong RCE demonstrates that some aspects of this extant body of work do not consistently align with the tenets of CRR.

Augmenting the cultural responsiveness of research underscores the potential for improving scholarship to pointedly address the needs of diverse students who are the fastest growing population in schools but who have historically experienced disparities in educational opportunities and outcomes. Identification of evidenced-based practices generated through CRR is also key for practitioners experiencing a double bind: an inherent tension between the complex ecologies of real life experiences, opportunities, and challenges that historically marginalized youth face, and the use of evidence-based practices that are not developed based on the preferences and real life experiences of diverse learners whom they serve. Because our work represents preliminary efforts develop and examine criteria for CRR, next steps include a formal and expanded effort to document the rubric's validity and reliability.

Knowledge production to solve complex problems requires multifarious tools and approaches. By focusing our efforts on RCE, we do not intend to imply that other quantitative and qualitative research need not include such considerations. Generalizability and the present privileged position of RCE in our field, however, demand that we attend to questions about CRR so that the evidence we generate is sound. We see this rubric as an initial tool that can augment the production of evidence and relevant, meaningful, sustainable change in educational practice through research that is culturally responsive.

Authors' Note

Both authors have made equal contributions to content of this article.

Acknowledgments

The authors thank Alfredo J. Artiles, Alba A. Ortiz, and Sylvia Linan-Thompson for their feedback on the initial rubric for culturally responsive research. The authors also thank Jim Wollack, director, and Sonya Sedivy, assistant scientist, from the University of Wisconsin–Madison Testing and Evaluation Services, for feedback on the initial rubric design.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

References marked with an asterisk indicate studies included in the rubric application.

- American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications: American Educational Research Association. *Educational Researcher*, 35(6), 33–40.
- American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice, and organizational change for psychologists. *American Psychologist*, 58, 377–402.
- American Psychological Association. (2005). *APA guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations*. Washington, DC: Author. Retrieved from <http://www.apa.org/pi/oema/resources/policy/provider-guidelines.aspx>
- American Psychological Association. (2010). Ethical principles of psychologists and code of conduct. Retrieved from <http://www.apa.org/ethics/code/index.aspx>
- Artiles, A. J., Kozleski, E. B., Trent, S. C., Osher, D., & Ortiz, A. A. (2010). Justifying and explaining disproportionality, 1968–2008: A critique of underlying views of culture. *Exceptional Children*, 76, 279–290.
- Arzubiaga, A. E., Artiles, A. J., King, K. A., & Harris-Murri, N. (2008). Beyond research on cultural minorities: Challenges and implications of research as situated cultural practice. *Exceptional Children*, 74, 309–327.
- Ashing-Giwa, K. T. (2005). Can a culturally responsive model for research design bring us closer to addressing participation disparities? Lessons learned from cancer survivorship studies. *Ethnicity & Disease*, 15, 130–137.
- August, D., & Hakuta, K. (1997). Priorities for research. In D. August & K. Hakuta (Eds.), *Improving schooling for*

- language-minority children: A research agenda* (pp. 343-362). Washington, DC: National Academy Press.
- Bailey, D. B., Jr., Skinner, D., Rodriguez, P., Gut, D., & Correa, V. (1999). Awareness, use, and satisfaction with services for Latino parents of young children with disabilities. *Exceptional Children*, 65, 367-381.
- Bal, A. (2011). *Cultural psychology: Understanding mind in culture in a diverse world*. Pre-congress workshop conducted at the annual meeting of the European Congress of Psychology (ECP), Istanbul, Turkey.
- *Bates, P. E., Cuvo, T., Miner, C. A., & Korabel, C. A. (2001). Simulated and community-based instruction involving persons with mild and moderate mental retardation. *Research in Developmental Disabilities*, 22, 95-115.
- Bazeley, P. (2007). *Qualitative data analysis with NVIVO*. Thousand Oaks, CA: SAGE.
- Bernal, G., Bonilla, J., & Bellido, C. (1995). Ecological validity and cultural sensitivity for outcome research: Issues for the cultural adaptation and development of psychosocial treatments with Hispanics. *Journal of Abnormal Child Psychology*, 23, 67-82.
- Calamaro, C. J. (2008). Culture competence in research: Research design and subject recruitment. *Journal of Pediatric Health Care*, 22, 329-332.
- Chard, D. J., Ketterlin-Geller, L. R., Baker, S. K., Doabler, C., & Apichatabutra, C. (2009). Repeated reading interventions for students with learning disabilities: Status of the evidence. *Exceptional Children*, 75, 263-281.
- Chouinard, J. A., & Cousins, J. B. (2007). Culturally competent evaluation for Aboriginal communities: A review of the empirical literature. *Journal of Multidisciplinary Evaluation*, 4, 40-57.
- Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional Children*, 75, 365-383.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Harvard University Press.
- Figueroa, R. A., & Newsome, P. (2006). The diagnosis of LD in English learners: Is it nondiscriminatory? *Journal of Learning Disabilities*, 39, 206-214.
- Fine, M., Weis, L., Weseen, S., & Wong, L. (2003). For whom? Qualitative research, representations, and social responsibilities. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 107-131). Thousand Oaks, CA: SAGE.
- Gallego, M. A., Cole, M., & The Laboratory of Comparative Human Cognition (LCHC). (2001). Classroom cultures and cultures in the classroom. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 951-997). Washington, DC: American Educational Research Association.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C. R., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, 71, 149-164.
- Graham, S. (1992). "Most of the subjects were White and middle class": Trends in published research on African Americans in selected APA journals, 1970-1989. *American Psychologist*, 47, 629-639.
- Gudykunst, W. B., & Lee, C. M. (2003). *Cross-cultural and intercultural communication*. Thousand Oaks, CA: SAGE.
- Harry, B. (1992). Making sense of disability: Low-income, Puerto Rican parents' theories of the problem. *Exceptional Children*, 59, 27-40.
- *Izzo, M., Cartledge, G., Miller, L., Growick, B., & Rutkowski, S. (2000). Increasing employment earnings: Extended transition services that make a difference. *Career Development for Exceptional Individuals*, 23, 139-155.
- Kozleski, E. B., Engelbrecht, P., Hess, R., Swart, E., Eloff, I., Oswald, M., & Swati, J. (2008). Where differences matter: A cross-cultural analysis of family voice in special education. *Journal of Special Education*, 42, 26-35.
- Ladson-Billings, G., & Tate, W. F. (2006). *Education research in the public interest*. New York, NY: Teachers College Press.
- Lareau, A., & Horvat, E. M. (1999). Moments of social inclusion and exclusion race, class, and cultural capital in family-school relationships. *Sociology of Education*, 72, 37-53.
- *Martin, J. E., Van Dyke, J. L., Christensen, W. R., Greene, B. A., Gardner, J. E., & Lovett, D. L. (2006). Increasing student participation in their transition IEP meetings: Establishing the self-directed IEP as evidenced-based practice. *Exceptional Children*, 72, 299-316.
- Merriam, S. B., Johnson-Bailey, J., Lee, M. Y., Kee, Y., Ntseane, G., & Muhamad, M. (2001). Power and positionality: Negotiating insider/outsider status within and across cultures. *International Journal of Lifelong Education*, 20, 405-416.
- Montague, M., & Dietz, S. (2009). Evaluating the evidence base for cognitive strategy instruction and mathematical problem solving. *Exceptional Children*, 75, 285-302.
- National Center for the Dissemination of Disability Research. (1999). *Disability, diversity, and dissemination: A review of the literature on increasing the utilization of rehabilitation research among diverse consumer groups*. Austin, TX: Southwest Education Development Laboratory.
- National Research Council. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- National Research Council. (2005). *Advancing scientific research in education*. Washington, DC: National Academy Press.
- *Nelson, J. R., Smith, D. J., & Dodd, J. M. (1994). The effects of learning strategy instruction on the completion of job applications by students with learning disabilities. *Journal of Learning Disabilities*, 27, 104-110.
- Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children*, 71, 137-148.
- Ortiz, A. A., & Yates, J. R. (2010). Enhancing scientifically-based research for culturally and linguistically diverse learners. *Multiple Voices*, 11, 13-23.

- Phillips, D. C. (2006). Muddying the waters: The many purposes of educational inquiry. In C. F. Conrad & R. C. Serlin (Eds.), *The SAGE handbook for research in education: Engaging ideas and enriching inquiry* (pp. 7–21). Thousand Oaks, CA: SAGE.
- Pope-Davis, D. B., Liu, W. M., Toporek, R. L., & Brittan-Powell, C. S. (2001). What's missing from multicultural competency research: Review, introspection, and recommendations. *Cultural Diversity & Ethnic Minority Psychology*, 7, 121–138.
- Richards, L. (2005). *Handling qualitative data*. London, England: SAGE.
- Scheurich, J. J. (1997). *Research method in the postmodern*. London, England: Falmer Press.
- Schneider, B. L., & Keesler, V. A. (2007). School reform 2007: Transforming education into a scientific enterprise. *Annual Review of Sociology*, 33, 197–217.
- *Sinclair, M. F., Christenson, S. L., & Thurlow, M. L. (2005). Promoting school completion of urban secondary youth with emotional or behavioral disabilities. *Exceptional Children*, 71, 465–482.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15–21.
- Slavin, R. E. (2003). A reader's guide to scientifically based research. *Educational Leadership*, 60(5), 12–16.
- Solano-Flores, G. (2008). Who is given tests in what language by whom, when, and where? The need for probabilistic views of language in the testing of English language learners. *Educational Researcher*, 37(4), 189–199.
- Subedi, B. (2006). Theorizing a “halfie” researcher's identity in transnational fieldwork. *International Journal of Qualitative Studies in Education*, 19, 573–593.
- Test, D. W., Fowler, C. H., Richter, S. M., White, J., Mazzotti, V., Walker, A. R., & Kortering, L. (2009). Evidence-based practices in secondary transition. *Career Development for Exceptional Individuals*, 32, 115–128.
- Tillman, L. C. (2002). Culturally sensitive research approaches: An African-American perspective. *Educational Researcher*, 31(9), 3–12.
- *Van Reusen, A. K., & Bos, C. S. (1994). Facilitating student participation in individualized education programs through motivation strategy instruction. *Exceptional Children*, 60, 466–475.
- Wells, S. J., Merritt, L. M., & Briggs, H. E. (2009). Bias, racism and evidence-based practice: The case for more focused development of the child welfare evidence base. *Children and Youth Services Review*, 31, 1160–1171.