Multicultural education professional development: A review of the literature

Hillary Parkhouse
Virginia Commonwealth University, heparkhouse@vcu.edu

Virginia R. Massaro
Virginia Commonwealth University, massarovrr@alumni.vcu.edu

Chu Yi Lu Qiu
Virginia Commonwealth University, luqiucy@mymail.vcu.edu

Follow this and additional works at: https://scholarscompass.vcu.edu/tedu_pubs

Part of the Teacher Education and Professional Development Commons

Recommended Citation

Parkhouse, H., Lu, C. Y., & Massaro, V. R. (2019). Multicultural education professional development: A review of the literature. Review of Educational Research, 89(3), 416-458.

Abstract

When their teachers are well equipped to foster inclusive and equitable classrooms, students from marginalized communities show higher rates of academic achievement, motivation, self-confidence, and self-efficacy. However, many teachers complete preparation programs feeling underprepared to work in culturally diverse classrooms, making high-quality professional development in this area crucial. We undertook a meta-ethnographic, systematic literature review of 40 studies of multicultural-education-focused professional development programs in order to better understand the forms and features of such programs that contribute to teachers' self-efficacy and success in working with culturally diverse students. We found a small literature base with too much variation across types of programs studied and outcomes analyzed to draw conclusions about the factors that contribute to effectiveness. However, the extant literature does point to important questions and considerations for both providers and researchers of multicultural education professional development. One area for future research is how PD providers navigate tensions or challenges arising from resistance to discussions of diversity and equity. Another is locating the balance between providing specific knowledge about students' cultures and guarding against promoting stereotypes or broad generalizations. Researchers and PD developers should also pay close attention to their underlying theories related to both teacher learning and multicultural education.

Keywords: literature review, multicultural education, professional development, cultural diversity

Multicultural Education Professional Development: A Review of the Literature

As societies all over the world become increasingly diverse, educators are often at the forefront of efforts to promote more equal distributions of opportunity and representation across groups. Multicultural education, peace education, and other pedagogical approaches have been developed to help schools cultivate the type of inclusive and open-minded culture that, if spread, could contribute to more peaceful, cohesive, and democratic nations (Banks & Banks, 2004; Elmi, 2010; Harber & Sakade, 2009). Multicultural and intercultural education has also had a second, complementary goal: to reduce inequitable academic outcomes among groups. Approaches such as culturally relevant pedagogy (Ladson-Billings, 1995), culturally responsive teaching (CRT; Gay, 2002), and culturally sustaining pedagogy (Paris, 2012) all aim to empower marginalized students to achieve academically while also embracing their heritages in the face of institutional forces that would otherwise subtract their cultural resources (Valenzuela, 1999). Teacher education and professional development programs have increasingly advocated these pedagogies, and many school districts have supported them as well.

Despite these efforts, educational systems in many countries continue to be characterized by disparities across groups. In the United States, Black and Latinx[1] students graduate from high school at far lower rates than White and Asian students (National Center for Education Statistics, 2018), and there are racial/ethnic disparities in many other achievement measures, such as test scores, credits earned, and enrollment in honors courses (Dee & Penner, 2017; Zirkel, 2008). Within other countries, similar disparities exist between dominant and non-dominant groups. In

---

[1] Although the studies reviewed here generally used the terms Hispanic and Latina/o, we use Latinx throughout this paper to be inclusive of non-binary genders and to be consistent with the usage of contemporary multicultural education scholars.

New Zealand, Māori students experience disproportionately high suspension rates, low credit attainment, and over-representation in special education, compared to students of European descent (Savage, Hindle, Meyer, Hynds, Penetito, & Sleeter, 2011). In China, achievement gaps persist between students of Han background and those from ethnic minority groups (Parkhouse & Rong, 2016). These are just a few examples illustrating the ubiquitousness of struggles to equalize educational quality for students of all backgrounds.

These educational inequities highlight the need for teachers to have the knowledge, skills, dispositions, and supports needed to ensure success for *all* of their students. Unfortunately, many educators complete teacher preparation programs without sufficient preparation to work in culturally diverse settings (Hollins & Guzman, 2005; McAllister & Irvine, 2000). Teachers often lack awareness of cultural and linguistic influences on student learning, accept inequities as inevitable, or do not consider attending to cultural diversity to be part of their professional responsibilities (Lee, Luykx, Buxton, & Shaver, 2007). Therefore, educators need frequent and high-quality professional development in this area during their induction period as well as throughout their careers.

Professional development (PD) programs on multicultural education[2] (ME) are prevalent, but research on such programs is scarce. We were unable to find a literature review that synthesizes the research on such programs. As a result, PD designers may have difficulty locating relevant research to inform their curricular decision-making. Designers and facilitators need to be able to find research-based evidence on the forms and features of ME PD that have

---

[2] We use the term *multicultural education* for this review because it was the most commonly used term in the literature we reviewed. However, we also acknowledge that some authors use *intercultural education* to denote a stronger emphasis on "the multiplicity and fluidity of cultures and cultural identities and the impacts of cultural shifts during moments of cultural exchange with others" (Cloonan, Fox, Ohi, & Halse, 2017, p. 131).

the greatest impact on teacher learning and ultimately student outcomes. Hoping to meet this

need, we undertook a synthesis of ME PD research.

This systematic literature review addresses the following questions: *1. Among the various*

forms of ME PD that have been studied empirically, what patterns are evident in terms of

activities, aims, content, duration and other features? 2. What are the impacts of these various

forms of PD on teachers and students? To answer these questions, we performed a systematic

literature review of 40 peer-reviewed journal articles and one scholarly book that together

examined 33 unique PD programs. We included studies of all programs, in all countries, that

targeted PK12 school-based professionals including administrators, mental health professionals,

and teachers. Before presenting the methods and findings of the literature review, we first situate

the study in the literature on professional development as well as multicultural education.

Professional Development: Frameworks and Impacts

For this study, we adopted Little's (1987) broad definition of PD as "any activity that is

intended partly or primarily to prepare paid staff members for improved performance in present

or future roles in the school districts" (p. 491). This can include workshops, teacher

inquiry/action research, coaching/mentoring, co-teaching, lesson study, virtual modules,

simulations, conferences, summer institutes, or various combinations of these (Darling-

Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Desimone, 2009; Gulamhussein, 2013;

Zepeda, 2012). PD aims to help practitioners learn new ideas and skills, make better

professional judgments, and improve upon their teaching practice in order to have a greater

impact on their students (Darling-Hammond et al., 2009; Shaha, Glassett, Copas, & Huddleston,

2016; Zepeda, 2012). Most teachers participate in some form of PD each year, and most school

districts provide PD with the goal of refining teachers' practices in order to improve student

achievement (Darling-Hammond et al., 2009; Shaha et al., 2016). A study of five urban school

districts in the United States found that, on average, these districts spent 3.6% of their total

operating budget (or $4,380 per teacher) on PD during the school year studied (Miles, Odden,

Fermanich, & Archibald, 2004). However, these authors also found, as had prior studies, that

districts underestimated the amount they were investing in PD and that the management of these

resources was scattered across multiple departments that were typically not coordinating with

one another. The selection and facilitation of PD initiatives tends to vary greatly across schools

and regions, making it difficult to draw conclusions based on comparisons or syntheses (Yoon,

Duncan, Lee, Scarloss, & Shapley, 2007).

Most studies of the effectiveness of PD programs have examined impacts on teachers'

knowledge and skills, rather than on student achievement (Yoon et al., 2007). However there are

a few exceptions. One meta-analysis examining only those nine studies that met the What

Works Clearinghouse standards for experimental design found that teachers who receive

"substantial" PD, or an average of about 49 hours, "can boost their students' achievement by

about 21 percentile points" (Yoon et al., 2007, p.iii). In a 16-study meta-analysis investigating

the effect of PD for K-12 math and science teachers on student achievement, Blank and de las

Alas (2009) found a positive impact in mathematics achievement with an average effect size of

0.21 for studies using a pre-post design. They also identified the following program elements as

effective: a focus on both subject matter and pedagogical strategies for teaching it, a duration of

at least six months, and "follow-up reinforcement of learning, assistance with implementation,

and support for teachers from mentors and colleagues in their school" (p. 27).

Garet, Porter, Desimone, Birman, and Yoon (2001) investigated a national sample of over

1,000 math and science teachers to better understand the relationship between "best practices" of

professional development as described in the literature and teachers' self-reports on changes to their teaching practices. They found that content focus, active learning, coherence, and collective participation each had a positive influence on teachers' learning. Content focus refers to whether the PD is centered on subject matter and embedded in practice, coherence is the alignment between the PD and teachers' knowledge as well as school, district, and state policies, and collective participation refers to whether teachers participate with colleagues from their own school versus teachers from across many schools. The authors found that the type of PD, i.e., "traditional" (e.g., workshop or conference) versus "reform" (e.g., study group or network) — had an effect only indirectly through its effect on the duration of the PD.

In part due to results like this one, Desimone (2009) proposed to the research community that studies of PD effectiveness should focus more on the features (e.g., duration, content focus) than the form, type, or structure of a given PD (e.g., workshop, action research, mentoring). She further argued that enough empirical research exists to establish consensus that the following five features constitute a framework for studying the quality of PD: content focus, active learning, coherence, duration, and collective participation.

Seven years later, Mary Kennedy (2016) expanded upon Desimone's (2009) framework using a review only of those PD studies that examined student outcomes, controlled for teachers' motivation to learn, had a minimum study duration of one year, and followed teachers, rather than students, over time. Kennedy found that PDs of higher intensity and longer duration did not always produce greater gains. In addition, she suggested that researchers have thus far paid too little attention to the relationships between PD facilitators and their participants, as well as teachers' desire for concrete strategies with well-developed rationales—as opposed to prescriptive directives or general content knowledge with no guidance of how to apply it. She

also found that when teachers willingly participate and are treated as equals to researchers/facilitators, they are more motivated and active in the PD, even when the content was prescriptive. These results suggest that the five features identified by Desimone may not encompass all the possible qualities that significantly influence the effectiveness of programs. These two studies also mirror a larger disagreement in the field over the power of particular features. For example, while some researchers cite the importance of extended duration (Blank & de las Alas, 2009; Garet et al., 2001; Yoon et al., 2007), others have found this characteristic was of less significance relative to other factors (Timperley, Wilson, Barrar, & Fung, 2007).

Another domain of research has examined the policy contexts that support impactful PD opportunities. Desimone and colleagues (2002), using survey data from 363 U.S. school district PD coordinators, found that the district implementation strategies that most contributed to the quality of PD were: aligning PD to standards and assessment, involving teachers in planning the PD, and continuous improvement efforts. A review of international studies of PD found that the following characteristics of school leaders positively influenced PD effectiveness in improving student outcomes: active support of teachers in PD opportunities and participation in PD as learners themselves (Timperley et al., 2007).

We still have much to learn about the characteristics of effective PD programs and the mechanisms through which they are effective. The majority of studies have been small-scale, often relied on self-report, and have frequently been conducted by the PD facilitators themselves. Even for those features viewed as having a strong empirical base, such as longer duration and active participation, we still do not have guidance on what thresholds of time or active engagement are needed (Desimone, 2009). In addition, we need a better understanding of how teachers incorporate new ideas into their practice, how learning transfers from the PD contexts

into the classroom, and how teacher motivation mediates professional growth (M. Kennedy, 2016). In undertaking this review of ME PD we attended to those features that Desimone, Kennedy, and others have highlighted as well as others not yet well studied (e.g., facilitator relationship to participants, teacher reflection, and whether the PD was tailored to the specific student population). Although Garet et al. (2001) concluded that the structure or form of PD may not play a large role in enhancing outcomes, we analyzed this data as well in order to confirm or complicate that prior finding.

Multicultural Education Frameworks

Multicultural education is an overarching term for the various historical and contemporary reform efforts to create more inclusive and equitable schooling for all children (Banks & Banks, 2004). Although James Banks (2004) attended to issues of power and inequity in his foundational conceptualization, implementation of multicultural approaches often consists of tacking onto the curriculum a few minority figures or superficial aspects of diverse cultures. This has led to what some called a "heroes and holidays" approach (Lee, Menkart, & Okazawa-Rey, 2011), which neglects hierarchies of power that accompany cultural difference and promotes simplistic stereotypes of cultures and groups.

Culturally responsive teaching (Gay, 2002) and *culturally relevant pedagogy* (Ladson-Billings, 1995) are specific approaches to multicultural education that refocus attention on educational equity and raising students' sociopolitical consciousness so that they can disrupt the tendency of schools to reproduce social inequalities. The newest generation of scholars in the field have developed *culturally sustaining pedagogy* as a means of emphasizing that marginalized cultures need to be actively maintained—not just accommodated—and that cultures

are both fluid and complex, sometimes containing elements that need to be problematized (e.g., misogyny and homophobia; Paris & Alim, 2017).

The field of teacher education has long struggled with the question of how to shift teachers' thinking about culture and diversity. McAllister and Irvine (2000) argued that teacher educators might better attend to teachers' readiness to learn through considering process-oriented models of how individuals progress through stages of racial identity development and cross-cultural competence (e.g., Helms' model of racial identity development and Bennett's Developmental Model of Intercultural Sensitivity). McAllister and Irvine argued that the field tends to focus more on the content of teacher education curriculum than on the process through which teachers learn.

Method

According to Petticrew and Roberts (2008), "[a] systematic review is a method of critically appraising, summarizing, and attempting to reconcile the evidence on a particular problem" (p. 198). We began our systematic review by attempting to locate all relevant articles using the databases Academic & Education research, ERIC, and EBSCOhost. We used the search term combinations listed in Table 1. We sought only articles published in or after 2000, the year of the last comprehensive review of multicultural education professional development we were able to find (i.e., McAllister & Irvine, 2000). We concluded our search in December 2017. These parameters produced an initial yield of 1,602 studies, which we considered sufficiently large to further justify our choice of time parameters.

After evaluating the 1,602 abstracts, we created the following criteria for selecting publications that were appropriate for this review: (1) the study examined a professional development program on one or more topics related to cultural diversity, such as intercultural

competence, culturally relevant and responsive pedagogies, or multicultural education; (2) the

study used original qualitative and/or quantitative data; (3) the professional development in the

study was designed for in-service teachers or other school professionals in PK12 settings; and

finally, (4) the study reported the outcomes of the PD, such as its impact on participants and/or

student academic performance. We included studies from any country.

Using the above criteria, we eliminated several studies; for instance, we excluded survey

studies that asked teachers to evaluate any prior ME PD programs they had attended (e.g.,

Gándara, Maxwell-Jolly, & Driscoll, 2005), conceptual articles or reviews of research (e.g.,

Saint-Hilaire, 2014), and studies that examined non-PD related school interventions like the

LEAD program in Canadian schools (Miles & Bailey-McKenna, 2016). At this stage we also

searched through the reference lists and used Google Scholar to find additional relevant studies.

In some studies, the authors mentioned that their article was part of a larger study, in which case

we located other papers related to that study when possible. These included the Nurturing

Mathematics Dreamkeepers (Marshall, DeCuir-Gunby, & McCulloch, 2012; Marshall &

McCulloch, 2010; McCulloch & Marshall, 2011) and the Transformative Professional

Development (Johnson, 2011; Johnson & Fargo, 2014; Johnson & Marx, 2009). In cases where

we were not able to locate additional publications, we contacted the authors to inquire about any

further papers that should be included in our review. Additionally, we conducted a hand search

of several teacher education journals including *Journal of Teacher Education* and *Professional

Development in Education*. We used these complementary methods of searching for studies in

order to minimize the likelihood that we overlooked any relevant studies (Cooper, 2017).

After reading the abstracts and screening the articles, a total of 48 publications were

selected for likely inclusion in the review. We then conducted a second round of elimination on

these publications. Each researcher carefully read 14 to 15 articles and documented (a) the

method used in the study; (b) a brief description of the PD program; (c) the theoretical

framework or approach; and (d) the key findings of the study.

We then had several meetings to decide which of these publications met all of our

inclusion criteria. We eliminated 8 articles at this phase, resulting in a final list of 40 studies to

be included in the review. For example, we eliminated a study by Cloonan, Fox, Ohi, and Halse

(2017) on intercultural education in Australia because the article described teachers' construction

of autobiographical narratives in several different contexts, which did not seem to constitute a

PD *program* per se. Because several articles examined the same PD program, only 33 unique

programs are examined across these 40 articles. Due to the small quantity of studies found, we

did not use quality of research design as a factor for exclusion. We followed Cooper's (2017)

recommendation to include studies that offer some data or implications that permit us to draw

useful inferences, even if the rigor of the research is not particularly strong.

Analyzing and Synthesizing Studies

The variability in the research methods used and outcomes explored within each of these

studies make a formal meta-analysis impossible. The researchers who collected qualitative data

on teacher impacts did not examine the same phenomena: some reported on teachers'

perceptions of PD effectiveness while others documented classroom practices. In terms of

quantitative data, a few authors collected student test scores (Bishop, Berryman, Cavanagh, &

Teddy, 2009; Grimberg & Gummer, 2013; Johnson & Fargo, 2014; Powell, Cantrell, & Correll,

2016; Savage et al., 2011; Sleeter, 2011); however, each study examined different grade levels

and types of assessment. The majority of quantitative data consisted of survey results, and the

survey instruments also varied across studies. Moreover, each study examined a very different

form of PD. With some focusing on science teaching for American Indian students and others

cultural responsiveness with Māori students, for example, it was difficult to draw comparisons

across outcomes.

We also cannot draw inferences about the specific types or features of PD that render the

PD more effective for learning about diversity because most studies relied heavily on qualitative

descriptions of changes in behaviors that are difficult to compare. For example, teachers in one

study self-reported and were observed using culturally relevant strategies they implemented after

the PD (Powell et al., 2016); while, in another, teachers described how they *planned to* better

incorporate home languages and knowledge (Brown & Crippen, 2016). In this case we cannot

determine which of the two PDs studied produced greater changes.

Therefore, we decided to take a meta-ethnographic approach to comparing the studies

(Noblit & Hare, 1988; Urrieta & Noblit, 2018). Rather than aggregating or synthesizing the

findings only, we attempted to "translate full storylines" (Noblit, 2018, p. 37). In other words,

we did not reduce studies to their salient themes, but instead preserved the full contexts of each

study to gain a better picture of what inferences could be drawn from the study as a whole. For

example, the PD studied by Lee et al. (2007) did not appear to produce dramatic changes in the

participants. Rather than focus on this finding, we carefully analyzed the hypotheses the authors

posited for *why* the PD was not more impactful, synthesizing these suppositions with the findings

and implications of the other studies. In the end, Lee et al.'s conjectures about the failure of their

program were equally if not more helpful than other authors' positive findings for our goal of

identifying considerations for future ME PD developers. Therefore, a meta-ethnographic

approach helped us reach a deeper and more nuanced understanding of the landscape of ME PD

than would be achieved through other forms of meta-analysis.

Coding schemes and process. We created several tables to compare all studies on a number of dimensions, including the problem of practice addressed, duration, participants, PD developer/implementer, and a detailed description of all activities occurring as part of the PD. Another table displayed the characteristics of the research methods, the outcomes for teachers and students (if applicable), any limitations (stated as well as unstated), and implications. In creating the table of results, we recognized the difficulty of comparing outcomes across such disparate programs and methodologies. We instead focused our energies on comparing characteristics of the PDs. We do nevertheless report the outcomes of each study, not for purposes of comparison, but so that readers of this review can more easily locate evidence of the outcomes of various instantiations of ME PD.

Identifying model type. A. Kennedy (2005) identified 9 types of PD programs: *training*, *standards-based*, *deficit*, *award-bearing*, *cascade*, *coaching/mentoring, community of practice*, *action research, and transformative*. None of the studies reviewed here classified as *cascade* or *award-bearing*, and in some ways all could classify as *deficit* in that the PD was intended to address a perceived need among the participants. *Transformative* and *standards-based* seemed to better characterize features of a PD rather than its overarching structure, so we were left with *training, coaching/mentoring, community of practice (CoP),* and *action research* as the four labels we would use to categorize the PDs. We replaced *training* with *workshop* because A. Kennedy defined the former as technocratic and "'delivered' to a teacher by an 'expert'" (p. 237), whereas many of the workshop models described in these studies included teacher-driven elements and content that went beyond technocratic skills. We independently coded all studies for the type of structure used (i.e., workshop, coaching/mentoring, CoP, action research, or a

combination). We met regularly and resolved any conflicting codes through discussion until agreement among all researchers was achieved.

   Identifying content focus. We then turned to the additional features of effective PD identified in M. Kennedy's (2016) framework (described above). While Desimone (2009) argued that a subject-specific focus on content may be the most important feature, M. Kennedy found that "when programs offering content knowledge were successful, the content was subsumed under a broader goal, such as helping teachers learn to expose student thinking" (p. 971). For our purposes, this is promising because developing culturally relevant pedagogies could serve as that broader goal. We report content focus here more for the reader's edification than to suggest this feature distinguishes effective ME PDs from less effective ones. We cannot evaluate this claim due to the incomparability of the PD programs and studies.

   Additional features. M. Kennedy (2016) also found that intensity/duration, coaching, and collective participation were not as predictive of student success as whether teachers were intellectually engaged and given the tools to apply their new learning in their classrooms. When we coded for these elements, however, we found that almost all studies attended to the intellectual engagement of participants (interrater reliability [IRR] = 96%)[3] and provided strategies with a rationale, rather than merely prescribing practices or presenting general bodies of knowledge without guidance on how to apply the knowledge (IRR = 87%).

---

[3] We calculated interrater reliability (IRR) across the three researchers for the following 5 codes: *intellectually engaging*, *specific to student community*, *offered by experienced teachers/teacher educators*, *pedagogical approach (i.e. prescriptive, strategy, insight, or body of knowledge)*, and finally whether the PD was *mandatory or volunteer*. Our overall IRR was 91%. After calculating IRR we deliberated to come to a consensus on how to characterize each study in the final paper.

We coded for three additional features that seemed important for our particular review. We determined whether the PD discussed cultural diversity broadly, or focused on the particular groups of students that the participants worked with (e.g., Spanish-speaking English language learners; IRR=96%). Taking up M. Kennedy's (2016) recommendation, we also coded for whether the PD facilitator was a teacher or teacher educator with deep knowledge of the profession and problems of practice. Unfortunately, only one (Lee et al., 2007) study explicitly described the facilitators' qualifications for leading the PD and, as a result, we could not systematically compare studies on this characteristic. Finally, we coded for whether the PD was mandatory or voluntary, noting M. Kennedy's argument that studies of mandatory PD may produce lower impact given that voluntary participants are more likely to be motivated to learn. In many cases this was not explicitly stated, leading to a relatively lower IRR (82%).

Results and Interpretation

Overall, we found far more differences among the programs than similarities. As mentioned above, one commonality among all but two (specifically, Katz, Inan, Tyson, Dixson, & Kang, 2010; and Papanastasiou & Conway, 2002) was a focus on intellectually engaging teachers through having them reflect on, analyze, and often create instructional materials to enhance their culturally responsive practices. Most employed a theoretical framework related to multicultural education (e.g., culturally responsive teaching). Those that did not instead selected theories related to teacher learning, professional development, or self-efficacy more broadly (i.e., Messiou & Ainscow, 2015; Molle, 2013; Tucker et al., 2005). Future ME PD researchers may benefit from articulating both the underlying theory of multicultural education as well as teacher learning from which they are operating. We return to this point in the discussion.

Theoretical Approaches

This review builds on M. Kennedy's (2016) meta-analysis of PD programs in which she argued for focusing not only on design features, such as length, but also on underlying theories of action, such as the problem the PD is intended to address and the pedagogies used to address it. Whereas the studies she reviewed aimed at four problems of practice (i.e., portraying curricular content effectively, containing student behavior, enlisting student participation, and exposing student thinking), the studies reviewed here aim at a fifth problem of practice: providing more inclusive and equitable education in diverse classrooms.

Specifying "diversity". The majority of studies took place in the United States and therefore focused on educational equity for Latinx students (Colombo, 2007), American Indians (Grimberg & Gummer, 2013), or a group often referred to by the vague term, "culturally and linguistically diverse (CLD) students" (Brown & Crippen, 2016; Psalti, 2007). One program in New Zealand focused on reducing gaps between Māori and White students, while one in Greece focused on teachers' cultural competence in working with immigrant students in general with no specific reference to nationality/ethnicity. This omission is representative of a dividing line we identified among these studies as a whole: about half (15 out of 33 studies) tailored the PD program to the particular group(s) of students with which the participants were working, whereas the other half discussed cultural responsiveness in more general terms. The former were able to offer culturally specific guidance, such the knowledge, traditions, and values of the Māori (Bishop et al., 2009; Savage et al., 2011; Sleeter, 2011) or the particular American Indian tribes within the community (Grimberg & Gummer, 2013). The latter had to rely on more generic models of culturally relevant pedagogy or cultural competence. Figure 2 displays the proportion of programs that were tailored to a specific culture or group.

Pedagogies to foster teacher learning. M. Kennedy (2016) identified four methods PD programs used to facilitate teacher enactment of new ideas within their practice: prescription, strategy, insight, and body of knowledge. Our review found few programs that solely transmitted a broad body of knowledge—for example of a particular culture or of culturally responsive teaching theories and practices. Those that did include this element also cultivated insights or provided strategies for how teachers might use this knowledge. For instance, one program taught general lessons about Latinx culture followed by specific concrete strategies for working with Latinx families (Smith & Bahr, 2014). Based on the evaluations of these PD programs collected across many of the studies, teachers greatly appreciated the concrete strategies and, when they were lacking, often requested that future PD incorporate more of them (Colombo, 2007; Devereaux, Prater, Jackson, Heath, & Carter, 2010). Powell and colleagues' (2016) study suggested that providing knowledge without strategies can merely raise awareness of the need for instructional modification without equipping teachers with the skills to do so, which could actually lower self-efficacy for working with diverse students.

Divergence in conceptualizations of multicultural education. Most authors identified culturally responsive teaching (Gay, 2002) and/or culturally relevant pedagogy (Ladson-Billings, 1995) as the overarching framework of the PD they studied. However, we noted some significant inconsistencies in the ways in which these approaches were conceptualized and operationalized across studies. Both approaches have an explicit focus on equity and empowerment, not just inclusion and intercultural understanding. Some of the programs that were intended to develop culturally responsive teaching lacked the critical stance inherent in such work (Ladson-Billings, 1995), and thus neglected the issue of equity while potentially reinforcing essentialized understandings of cultures. For instance, the only example Johnson

(2011) provided of how teachers drew on students' cultural funds of knowledge was to use a

tortilla press to demonstrate a type of simple machine. Not enough details about the lesson were

provided to determine if this was done in a way that affirmed and engaged students or that may

have reinforced stereotypes about Latinx students. Another teacher in the study, who was

described as having grown in cultural responsiveness, was quoted as saying that if students do

not do well in this class, it is because "they choose not to do well" (Johnson, 2011, p. 181). This

statement reflects a lack of critical self-reflection and awareness of the many forces—beyond

motivation—that may impact a student's success in a class. The one example Brown and

Crippen (2016) used to illustrate how teachers used the cultural experiences of students to

facilitate learning was a lesson on racial differences in skin cancer and genetic disorders, such as

Tay-Sachs and sickle-cell anemia. Depending upon how this lesson was implemented, it could

have had the opposite of the intended effect of culturally responsive teaching by perpetuating the

myth that race is a biologically, rather than socially, determined difference in humans (Smedley

& Smedley, 2005).

We also noted that some authors conflated culturally responsive teaching with generally

effective pedagogy by defining the former as scaffolding, using a variety of formative

assessments, pointing out misconceptions, and building lessons on prior learning. By contrast,

Geneva Gay's (2002) definition of culturally responsive teaching (CRT) is "using the cultural

characteristics, experiences, and perspectives of ethnically diverse students as conduits for

teaching them more effectively" (p. 106). Authors describing the Transformative Professional

Development (TPD) program argued that the two focal teachers of the study were using

culturally relevant strategies like cooperative learning, differentiation, inquiry, and setting high

expectations (Johnson, 2011). These are indeed elements of CRT, but without links to students'

cultural backgrounds, these strategies could also be used in non-culturally relevant or responsive ways. Our analysis reveals that some reported positive impacts of PD programs may be grounded in loose interpretations of CRT (see also Zozakiewicz & Rodriguez, 2007).

Components of the Programs

Most (29) of the PD programs involved workshops; although 17 of these incorporated one or more additional experiences, such as coaching (5), action research (12), or communities of practice (10). Figure 1 displays the wide range of experiences involved in each program. Each single row within each column represents one of the 33 professional development programs included in this review. Reading across the rows of the four-component column, for example, allows one to see that, of those 7 studies, two consisted of *workshop* plus *action* plus *community of practice* plus *coaching*. A third study was similar except its fourth element was *video feedback* rather than *coaching*, and so on.

Given that discussions of race, language, and identity can evoke cognitive dissonance, feelings of guilt or defensiveness, and questions about self-efficacy (Wing Sue, Torino, Capodilupo, Rivera, & Lin, 2009), it may be that designers of this type of PD are especially interested in structures that will help participants take ownership over their own learning (such as action research), both to build efficacy and to allow each person to progress at their own pace. Some teachers may need to try out elements of culturally responsive education themselves before they are convinced of its relevance to their subject area and their own capacity to enact it. In addition to coaching, communities of practice, and action research, three unique structures incorporated into a few of the programs reviewed were "critical friendships" (Johnson & Marx, 2009, p. 125), self-evaluations of teaching, and field experiences, such as visits to homes,

religious centers, or events for parents. The next two sections describe these structures in more

detail.

Critical friendships. Both the Nurturing Mathematics Dreamkeepers program (NMD;

Marshall et al., 2012; Marshall & McCulloch, 2010; McCulloch & Marshall, 2011) and the TPD

program (Johnson, 2011; Johnson & Fargo, 2014; Johnson & Marx, 2009) set up critical

friendships in which teachers observed peers' lessons and gave constructive feedback. In the

NMD program, participants were assigned same-grade "buddies." Buddies observed each other

teaching two consecutive math lessons via video-recordings four times per year. After each

observation, the buddies engaged in a structured reflective session in which they discussed the

students' mathematics reasoning and the role of culture in the lesson (Marshall et al., 2012).

This system was innovative in that participants served as both mentees and mentors to one

another. The drawback of this model is that the coach may have limited expertise; however,

serving as a coach may bolster each participants' self-efficacy and deeper knowledge of

culturally relevant education, in which case the advantages may outweigh the disadvantages (see

also Savage et al., 2011).

Within the TPD (Johnson & Marx, 2009) program, teachers were provided with

substitute teachers so they could observe each other's classrooms and give constructive feedback

within these newly formed "critical friendships" (p. 125). The authors explained that this

allowed the teachers to observe both effective practices they could adopt as well as ineffective

practices that they might recognize as methods they use themselves. However, this component

of the multi-faceted TPD program was not described in detail nor mentioned in the authors' other

articles (i.e., Johnson, 2011; Johnson & Fargo, 2014), suggesting that it was not a major

component of the program. The workshops on inquiry-based science and culturally relevant

pedagogy seemed to be the central elements. Nevertheless, the authors implied that these critical

friendships, in combination with monthly focus group discussions about their instruction,

contributed to the following result: "Collective teacher leadership and empowerment enabled

individual classrooms and climates of schools to begin to transform" (Johnson & Marx, 2009, p.

129).

      Self-evaluation through rubric rating and video analysis. In several programs,

participants provided critical feedback to themselves rather than to peers, through using rubrics

to score themselves or through viewing videos of their own classrooms. In Hulan's (2015) study

of a book club, teachers were videotaped while teaching, and the researcher and teacher viewed

it together to discuss discourse practices as studied in the book club. In Trifiro's (2017) study of

a PD to support teachers' instruction of English language learners (ELLs), participants scored

themselves on the eight instructional practices designed for ELLs before and after the program

(Echevarria, Vogt, & Short, 2016). In final surveys, participants rated this activity as one of the

three most valuable learning activities of the whole one-year program (the other two were not

listed).

      Within the NMD program described above, teachers in pairs viewed videos of their own

and their buddy's instruction, and the two analyzed the practices of both (Marshall et al., 2012).

The authors also independently coded these videos for culturally relevant teacher-student

interactions and teachers' efforts to highlight or clarify mathematics thinking. They found that

teachers' interpretations of their instruction sometimes differed markedly from the researchers'

analysis. For instance, one teacher, Aisha, described the importance of meeting children where

they are and connecting to their outside-of-school experiences, but "her classroom practices

reflected obliviousness if not disregard for children's prior mathematical knowledge" (Marshall

et al., 2012, p. 12). The authors viewed this as evidence that the high-stakes testing environment

and teacher's lack of comfort with the mathematics curriculum made it difficult for Aisha to

embody those culturally relevant practices she valued. The researchers admitted that they had

erroneously assumed that Aisha would demonstrate culturally relevant pedagogy because she

expressed racial consciousness and sociopolitical awareness as an African American woman.

This study highlights the fact that dispositions by themselves may not translate to culturally

responsive practices, particularly in high-stakes testing environments or in cases in which the

teacher is still developing content knowledge.

Immersion experiences. Several programs utilized home visits so that participants could

gain a deeper understanding of their students' cultures and community funds of knowledge

(Johnson, 2011; Devereaux et al., 2010; Prater, Wilder, & Dyches, 2008). The PD for education

faculty studied by Prater et al. (2008) had participants conduct home visits to their special

education and ESL teacher candidates' homes to interview them about their personal journeys,

successes and challenges, and recommendations for improvement of the program. The TPD

program described above also incorporated home visits, as well as a conversational Spanish

language course, in part because many of the teachers did not live in the school community and

lacked awareness of community assets. The Johnson study of the impacts of TPD on two

teachers found that the home visits helped them to realize that the parents of their Latinx

students did value education and support their children, contrary to their beliefs prior to the PD.

However, they also found that the experience did not prompt the two teachers to become more

involved in the community around their school.

Two other programs had innovative ideas for other types of cultural immersion

experiences beyond home visits. In the Parent Partnership for Achieving Literacy (PAL),

teachers attended two of the family literacy nights that were offered twice weekly throughout the year. These events consisted of parent-child reading activities, homework help for students, and ESL classes for parents (Colombo, 2007). Teachers attended in groups no larger than two, in order to ensure a more immersive experience and intensive interactions with the families. They participated in reading, singing, and playing games in Spanish as well as assisting with the ESL classes. Twenty-five of the 27 teachers interviewed found the experience useful, explaining that "feeling lost" during these literacy nights helped them to put themselves in their ELL students' shoes and to recognize the hegemony of English in U.S. schools (Colombo, 2007, p. 13). It also helped participants recognize some of their misconceptions about Latinx culture and cultivate an appreciation for bilingualism. However, for two teachers (who had attended the nights together), the literacy nights reinforced deficit perspectives they held about Latinx parents and their capacities for helping their children. The researchers recommended that future programs ensure such teachers are regrouped so that their perspectives are challenged rather than reinforced by their peers.

The other cultural immersion experience studied was a trip to a local Gurdwara (Sikh temple) and mosque (Islamic center) as part of a PD for a school in Calgary, Canada in which most of the students were English language learners, primarily from South Asian backgrounds (Aujla-Bhullar, 2011). Participants admitted they never would have visited these sites "by choice" (p. 271) and that they were surprised by how welcoming the centers were. They reported that the visits helped them better understand the religions and role of religion in their students' lives as well as appreciate the importance of maintaining those beliefs and traditions. However, the researcher (also PD facilitator) noted that a "silencing effect" was produced when participants hesitated to share certain thoughts and instead "pushed [them] away to maintain a

'positive' environment" (p. 272). She recommended that PD facilitators be sensitive not only to the explicit needs of participants but also the submerged needs that may arise from their values or ideological beliefs. No specific recommendations for how to do so were shared. Readers interested in immersion as a form of PD for multicultural education may wish to consult the extensive body of literature on study abroad and teaching abroad as a disorienting and transformative experience for teachers of all backgrounds (e.g., Jiang & DeVillar, 2011; Marx & Moss, 2011; Parkhouse, Turner, Konle, & Rong, 2016; Quezada, 2004; Trilokekar & Kukar, 2011).

Duration and Intensity

The duration of PD activities ranged from a one-day workshop (Aujla-Bhullar, 2011; Smith & Bahr, 2014) to five years of monthly trainings along with a two-week summer institute (Grimberg & Gummer, 2013). We cannot determine a relationship between duration and effectiveness due to the difficulty of making side-by-side comparisons of these studies (described in the method section). We can say, however, that even the briefest programs reported positive impacts evident at least three-to-nine months later (Charity Hudley & Malinson, 2017; Smith & Bahr, 2014). On the other hand, the participants in Lee and colleagues' (2007) study showed little growth despite the fact that their program spanned four full days per year for two years. These findings lend support to M. Kennedy's (2016) assertion that the programs with shorter durations may not necessarily be the least impactful. This is not to suggest that duration makes no difference; however, we agree with M. Kennedy's argument that focusing primarily on the length or intensity of a program may draw attention away from potentially more important characteristics, such as whether the program is intellectually engaging and meaningful to participants.

Overall, none of the studies attempted to specifically measure the significance of the duration, frequency, or intensity of their PD program. However, Hulan (2015), who studied her book club on multicultural literacy, concluded that meeting every two weeks contributed to teacher camaraderie, allowed the material to stay fresh in their minds, and ensured they were able to get feedback quickly. Four groups of researchers received feedback from teacher participants that the time they received was inadequate and more was needed to unpack the complex ideas presented (Aujla-Bhullar, 2011; Devereaux et al., 2010; Mette, Nieuwenhuizen, & Hvidston, 2016; Timmons-Brown & Warner, 2016).

Audience

As shown in Figure 3, the majority (21) of the PD programs studied were not geared toward a specific content area; however, four were designed for science teachers, four for math teachers, two for STEM, and one for social studies. No studies targeted English teachers, although one was geared toward literacy instruction for elementary teachers (Hulan, 2015). In addition, one program was designed for special education faculty who work with pre-service teachers (Devereaux et al., 2010; Prater et al., 2008). Figure 4 depicts the proportion of programs that were aimed at educators at the PreK, elementary, middle, high, and higher education levels.

One study targeted school psychologists and social workers, rather than teachers (Smith & Bahr, 2014). Those authors found no other studies on PD for this audience, suggesting that this is an important area for future research. The only study we found that targeted school administrators examined a racial autobiography project prospective administrators completed as part of their university coursework (Gooden & O'Doherty, 2015). We did not consider this to be a professional development program and thus did not include it in our review.

Figure 5 illustrates the proportion of programs that were mandatory versus voluntary. Many authors did not explicitly state this information, but we inferred that all programs that were provided to an entire department, school, or school division were probably mandatory. However, because it is possible that an entire department or school may have elected to participate, we coded such programs as "all staff."

Facilitator

M. Kennedy (2016) found that more effective PD programs were offered by people who had "had long histories of working with teachers, were very familiar with teachers and with the problems they face, and based their programs on their own personal experience and expertise" (p. 973). This seemed particularly important for PD related to sensitive topics, such as cultural diversity and inequities in schools, in which the success of participants who are relatively less familiar or open to these discussions may be largely dependent upon a trusting relationship between the participants and PD facilitator. Unfortunately, most of the studies here provided insufficient details about the facilitator to adequately evaluate this component. However, 24 of the 33 seemed to have been facilitated by the authors, who were professors in colleges of education and thus likely had some history of working with teachers and may have been teachers themselves earlier in their career. Hulan (2015) was one of the few authors to comment on the relationship between the facilitator and participants. She made frequent visits to each of her participants' schools, becoming "more of a co-participant than an outsider," which "helped to build trust and to make the professional development more successful" (p. 72). Our review supports M. Kennedy's assertion that this is an element of PD that deserves greater attention.

Impacts on Participants and Students

In our above descriptions of the forms and features of PD programs, we included a few of the impacts and limitations of each PD program on participants. This section will provide some of the broader themes we identified related to impacts. All 33 of the programs measured impacts on teachers, with only six additionally measuring impacts on students (detailed below). Of those focusing solely on teacher impacts (28 programs), 3 reported growth through quantitative measures, 18 through qualitative evidence, and 7 through a mixed-method design. Very few used a logic model or other means of identifying the specific intervening variables that contributed to the positive outcomes (e.g., teacher knowledge). All but one of these studies (i.e., Lee et al., 2007) deemed the PD program to have been a success based on their data. Lee and colleagues (2007), however, offered some of the most helpful analysis in terms of our synthesis of this body of knowledge (detailed below); therefore, we want to underscore the importance of publishing studies of programs that fail to achieve the results anticipated, as these can be as informative as the programs that fulfill their goals.

Outcomes for teachers. The studies demonstrated a wide range of benefits to teachers based on data from interviews, surveys, and classroom observations. Nine studies relied only on questionnaires, three relied only or primarily on interviews, eight used both questionnaires and interviews, and 10 added classroom observations to their interview and/or survey data. See Table 2 for a list of the impacts of the various programs on the participants.

Some authors indicated that the PD participants, in some ways, did not grow as much as they had initially hoped or expected (Aujla-Bhullar, 2011; Brown & Crippen, 2016; Lee et al., 2007; Watkins & Noble, 2016). For instance, Brown and Crippen (2016) wrote the following regarding the participants in their program: "While we documented evidence of teachers connecting science content with students' racial backgrounds (Adriana) and out-of-school

experiences (Laura), a more-pronounced eliciting of students' and families' funds of knowledge would have been desirable" (p. 488). They determined the main problem to be that teachers possessed limited knowledge about their students' cultures.

Several studies found that some participants avoided or resisted fully engaging in these difficult conversations (Aujla-Bhullar, 2011; Colombo, 2007; Johnson & Fargo, 2014; McCulloch & Marshall, 2011; Prater et al., 2008; Watkins & Noble, 2016; Wiseman & Fox, 2012; Zozakiewicz & Rodriguez, 2007). Researchers of the NMD program found that when teachers were not drawing superficial connections between race and culture, they were often taking a colorblind perspective out of discomfort with explicit discussions of race (McCulloch & Marshall, 2011; see Zozakiewicz & Rodriguez, 2007 and Wiseman & Fox, 2012 for additional examples of resistance and denial as responses to multicultural PD).

Some teachers believed their subject matter (e.g., mathematics) was objective and therefore not related to issues of cultural diversity (Timmons-Brown & Warner, 2016), while other teachers believed they were already practicing multicultural education and did not need to change (Zozakiewicz & Rodriguez, 2007). More veteran teachers were sometimes skeptical of PD that asked them to modify their instruction, as their methods to this point seemed to have been effective (Johnson & Fargo, 2014; Zozakiewicz, & Rodriguez, 2007). As M. Kennedy (2016) pointed out, PD often asks teachers not only to adopt a new approach but also abandon an old one, and as a result, veteran teachers may find this particularly challenging or aversive. However, one study suggested teachers may be motivated by a conversation that opens with a recognition of their existing strengths (Smith & Bahr, 2014). Another source of resistance could be unfamiliarity with the relationship between multicultural education and academic outcomes. As one participant stated, "I know universities are driving this whole multicultural thing but

there is some[thing] else driving our school—academic results! . . . teachers need to be in the classroom teaching" (Watkins & Noble, 2016, p. 52). Recently, two large-scale studies have shown links between culturally relevant education and GPAs as well as attendance rates of students deemed at risk of dropping out of school (Cabrera, Milem, Jaquette, & Marx, 2014; Dee & Penner, 2017).

Several studies reported a lack of growth among participants at the lower ends of the cultural competence continuum (Colombo, 2007; Lee et al., 2007; Rubel & Chu, 2012) or reported that respondents believed many of their colleagues continued to show apathy in regard to cultural diversity (Mette et al., 2016). Lee and colleagues (2007) were the only authors to offer a comprehensive hypothesis for why their PD was not more effective. They posited: (a) the teachers resisted infusing culture into the curriculum, (b) the intervention raised awareness but did not change behaviors, (c) teachers lacked materials that connected students' cultures to the curriculum, (d) the PD was mandatory and thus some participants were not motivated to learn, (e) the policy context of high-stakes assessment and accountability created major challenges to the intervention, and (f) the state in which the PD occurred required students to be moved into English-Only instruction as quickly as possible, thereby discouraging teachers from developing or maintaining home languages. These challenges for ME PD will be further discussed in the final section.

Outcomes for students. Of the six studies that included student outcomes, two found qualitative improvements in students' written work and feelings of connectedness to school (García & García, 2016; Sleeter, 2011), three found improvements in test scores (Johnson & Fargo, 2014; Mette et al., 2016; Powell et al., 2016), and one found no impact on test scores but did find that teachers showed increased confidence in teaching culturally diverse students

(Grimberg & Gummer, 2013). This last result is surprising, given that Grimberg and Gummer's program was intensive and carefully designed. Their three-year PD program involved leaders of local American Indian tribes in designing curriculum and facilitating PD to integrate students' American Indian cultures into science instruction. Although teacher survey results revealed growth in teachers' confidence in their abilities to teach equitably and to prompt students to make connections between science and their own lives, the treatment groups did not show greater gains on science test scores than the control groups did.

However, three studies did find that their PD program produced a growth in test scores. Johnson and Fargo (2014) measured changes in elementary students' scores on state science exams after their teachers participated in the two-year TPD program (described above and in the appendix) and compared these with students in a school that did not receive the program. They found that proficiency rates rose from 25% to 67% in the treatment school over the two years, whereas the control school rates went from 18% to 29%. They also found a significant effect of the treatment condition for Latinx students but not non-Latinx students.

Mette et al. (2016) compared non-White students' ELA, mathematics, and science test scores across two years: the first year of the PD sessions and the following year (during which teachers were still participating in the PD). They found the proficiency rate in ELA rose from 40.1% to 52.2% and mathematics 8.5% to 35.5%, but science decreased from 64.5% to 48.3%. Mathematics proficiency rates for non-White students were also higher in the treatment school than across the district following the PD.

Finally, Powell et al. (2016) found that the students of those PD participants that were scored as showing high implementation of culturally relevant instruction (CRI) produced significantly higher achievement scores in reading and mathematics than students of teachers

with low implementation of CRI. Given the small number of studies that have examined impacts on students, more research is needed to better understand the precise ways in which multicultural education-related PD does and does not affect student academic outcomes.

A nuanced understanding of student outcomes might include data other than standardized test scores, such as student surveys, interviews, or classroom artifacts. The researchers studying Te Kotahitanga (a PD program developed in collaboration with Māori students and for Māori students) used such data in addition to quantitative measures including but not limited to test scores. For instance, they measured retention rates over three years and found that, in schools implementing Te Kotahitanga, there was a 260% increase in retention past grade 11, which is the grade after which schooling is no longer mandatory in New Zealand (Sleeter, 2011). In interviews, students expressed that they felt more respected and appreciated that their teachers spoke Māori to them, attended community events, and related lessons to Māori knowledge (Savage et al., 2011). One important insight that emerged from this study is that care must be taken when developing such programs to ensure students do not think the purpose is to remediate them. If they had not interviewed students, the researchers would likely never have known this was one perception of the program among some Māori students.

García and García (2016) used three sets of data to measure the impact of a teacher action research project on the students in two heritage Spanish classes: writing samples, interviews, and a survey that asked students about the strength of their ethnic identification. Comparison of writing samples before and after the project revealed progress in students' content and organization, and survey results showed increases in cultural pride and positive regard toward members of their own ethnic group. Seeing this growth inspired the focal teacher to develop school-sponsored, communitywide events to promote bilingualism and biculturalism.

Discussion

Recognizing the nascent state of the empirical research on ME PD, we offer this review not as a means of identifying the best practices for such PD but rather to organize the existing knowledge in one place so that PD developers and researchers can easily find the types of program designs that have been tried and the impacts associated with various forms of PD. An additional upshot of this literature review is the synthesis of areas in which the PD programs did *not* achieve desired results, along with identification of areas most in need of research. PD researchers may design studies to better understand and address those areas that appeared to be associated with limited impacts. In the remainder of this article, therefore, we detail the two domains that appear, from this review, to be most important to consider: the challenging nature of this work and lingering questions about how teachers learn and change behaviors. We also recommend methodological approaches in future ME PD research that may facilitate comparison across studies and thus a stronger empirical grounding for ME PD development.

Challenges Unique to Multicultural Education Professional Development

Many of the challenges faced by the PD providers in these studies are common across all types of PD, whether focused on multicultural education or not. For instance, programs of all types have to compete against the many other initiatives faced by teachers and have to be compelling enough that teachers are willing to change their current practices (Desimone, 2009; M. Kennedy, 2016). However, this review has highlighted a few of the challenges that are unique to professional development related to multicultural education.

Silence, resistance, and awareness of social structures. Discussions of racial, socioeconomic, religious, and other forms of cultural difference are often deeply personal and emotionally taxing—or avoided altogether due to their political nature (McAllister & Irvine,

2000). In the studies reviewed here, participants were prompted to reflect on their own identities, their attitudes toward difference, and their responses to entrenched societal problems, such as racism, sexism, linguicism, and ableism. In several cases these conversations produced silence or resistance, or in other ways failed to transform teaching practices (Aujla-Bhullar, 2011; Brown & Crippen, 2016; Colombo, 2007; Watkins & Noble, 2016; Wiseman & Fox, 2012). Only one study (i.e., Molle, 2013) closely analyzed how the PD facilitator responded to tensions arising from ideological differences; however that study did not examine the perspectives or practices of participants afterward. Teacher education scholars have recommended attending to affective dimensions of professional learning, in addition to cognitive dimensions, and being prepared for "complex and contradictory emotional reactions" so one can "engage them as moments of pedagogical opportunity in and of themselves" (Garrett & Segall, 2013, p. 302). Future ME PD research could make a tremendous contribution to the teacher education field as a whole by closely investigating facilitators' responses to instances of resistance or silence, as well as how these responses influence outcomes for participants.

Both cognitive and affective or dispositional shifts may be necessary to embrace culturally relevant pedagogy and its underlying proposition that social structures shape individual behaviors as well as institutions (Ladson-Billings, 1995). Without an awareness of social structures, discrimination of all forms tends to be viewed as residing solely within individuals rather than also in institutions and cultural practices. However, for many teachers, interpreting social phenomena through a structural lens may require much more than a single PD program, particularly if the program focuses on curriculum design in addition to perspective shifts. In this review, the programs that focused on content-specific strategies for enhancing culturally responsive instruction appeared not to save sufficient time for fostering this broader

understanding of society. This is not to suggest that a single PD could not enhance both content-specific cultural responsiveness as well as a structural understanding of inequality—in fact, M. Kennedy's (2016) meta-analysis suggested that content-specific PD may result in greater gains for students if the content strategies are subsumed under a broader goal. That broader goal could be inclusive and equitable instruction, for example. The point is that meeting both goals likely requires a period of time that is not only prolonged but also protected from the multiple other demands on teachers (Hulan, 2015; Johnson, 2011; McCulloch & Marshall, 2011). We need more research into how perspective shifts relate to interactions between program intensity, content specificity, and other factors, such as policy contexts and backgrounds of participants.

Studies of schools that have been successful with culturally diverse populations suggest a whole-school approach can help create a schoolwide culture of self-reflection and openness to new approaches and ways of thinking (Mayfield & Garrison-Wade, 2015). More research is needed to identify the types of school contexts and professional cultures that may help teachers develop and act on deeper understandings of diversity and inequality. In addition, it would be helpful to know more about the practices that PD providers can use to build trust and create an atmosphere encouraging of self-reflection, vulnerability, and openness. Such knowledge would strengthen capacities for ME PD to minimize resistance and silence among participants and encourage structural awareness and a commitment to lifelong growth.

Balancing specific information with avoidance of stereotypes. The other challenge revealed through this literature review is navigating the fine line between ensuring a PD opportunity offers specific information about responsive teaching for particular student groups and not reducing these groups to stereotypes or simplistic characterizations (see Schmeichel, 2012 for an excellent exploration of this tension inherent in culturally relevant pedagogy). A few

of the examples of culturally responsive teaching provided in these studies seemed at risk of

promoting essentialized perceptions (e.g., Brown & Crippen, 2016; Johnson, 2011), while others

were so broad that they appeared not to draw on any particular cultural assets or funds of

knowledge (e.g., Katz et al., 2010; Marshall et al., 2012). To address the first risk, some

programs reviewed involved members of the local community. For instance, the Grimberg and

Gummer (2013) PD involved local tribal leaders, and the Te Kotahitanga PD providers in New

Zealand sought input from the Māori students in order to design the program (Bishop et al.,

2009; Savage et al., 2011; Sleeter, 2011).

More research is needed to help both PD developers and participants locate the fine line

between specificity and reduction to achieve nuanced, complex understandings of cultural

influences on student learning. We agree with Watkins and Noble (2016) when they argued that

a strong multicultural education program "must approach these elements as a fundamentally

intellectual task, not simply an ethical one of tolerance of cultural difference" (p. 54, emphasis

original). Honoring the intellectually complex nature of culturally responsive teaching could

help educators foster the critical, nuanced understandings of diversity and equity necessary for

thinking through how social conditions relate to their specific classroom contexts.

Lingering Questions about How Teachers Change

In addition to these two challenges unique to ME PD, researchers and PD designers may

also benefit from humility and reflectiveness regarding the degree to which we currently

understand how teachers change their practice. Within the path model Desimone (2009) created

as part of her core conceptual framework for studying the effects of PD, changes in teaching

practices come about after attitudes, knowledge, and skills are enhanced. Although not explicitly

stated in the studies we reviewed, most seemed to be operating under the same logic or, as M.

Kennedy (2016) called it, "PD theory of action" (p. 947). Few explored the possibility that teachers might alter practices without altering their underlying beliefs (see DeJaeghere & Cao, 2009 for one exception). Most also neglected what M. Kennedy called *the problem of enactment* which is "a phenomenon in which teachers can learn and espouse one idea, yet continue enacting a different idea, out of habit, without even noticing the contradiction" (p. 947).

A few authors did mention cases in which participants demonstrated this problem of enactment, such as Aisha in Marshall et al.'s (2012) study (described above). The case of Aisha suggested that teachers need strong pedagogical content knowledge as well as an environment conducive to such pedagogies in order to effectively connect classroom learning to their students' out-of-school experiences. Regarding the use of video for professional growth, the study also suggested that showing teachers videos of their own instruction may not be enough to alert them to the areas in which they could improve.

In other cases, teachers did not demonstrate the problem of enactment but did show a similar difficulty in evaluating the effectiveness of their culturally responsive practices. Mette et al. (2016) found that teachers perceived their PD as having relatively little impact on the achievement gap between White students and students of color at their school, when in reality the gap shrunk the year following the study. Taken together, these findings regarding the problem of enactment and other reality-perception mismatches suggest that teachers may need additional supports in order to be able to accurately assess their efficacy. Future ME PD research should explore cases that illuminate the reliability of the path model described above and what deviating cases can tell us about the range of ways in which teachers grow professionally. PD providers may want to consider Marshall et al.'s (2012) recommendation to have teachers not only reflect but also "unpack contradictions between their professed beliefs,

their teaching practices, and their explanations for those teaching practices" (p. 17).

Methodological Recommendations for Future Research

This literature review revealed that a great deal of additional research is needed to further our understanding of the forms and features of ME PD that most contribute to practitioner growth. Not all of the studies here collected baseline data for the teachers or students, and thus strong claims could not be made about changes resulting from the program studied. We also hope to see more large-scale studies as well as small-scale studies that at least partially replicate prior studies or use the same measures so that results can be compared and synthesized. Given the general paucity of qualitative, quantitative, and especially mixed methods research on ME PD, we are less concerned with which data collection methods are used (more of all would be helpful), and more concerned that researchers build on prior work and attend to the degree to which their results can be compared to others'. To borrow from meta-ethnographers' terminology, greater *translatability* (Urrieta & Noblit, 2018) across studies would provide warrant for the types of specific recommendations that we are not able to make given the current status of this body of literature.

For measuring student growth, triangulating across multiple data sources would be ideal, given the limitations of using student test scores alone as a measure for teacher performance (Baker et al., 2010; Stronge, Ward & Grant, 2011). The researchers of the Te Kotahitanga program demonstrated how student impacts can be more thoroughly and compellingly analyzed through supplementing traditional measures of achievement with measures of student retention and work completion, along with focus group data (Bishop et al., 2009; Savage et al., 2011; Sleeter, 2011). The focus groups illuminated students' feelings towards school and perceptions of how much their teachers cared about them and understood Māori culture. Although test

scores can be influenced by a number of factors beyond teaching practices, students' feelings about those practices are generally influenced only by the practices themselves and the students' preferences. Additionally, qualitative data such as these may help reveal not only *whether* PD-influenced changes in instruction affected student outcomes but also *how* they did so (e.g., through helping youth feel more connected to school).

In order to identify which particular components and characteristics of ME PD contribute to greater impact, it would be helpful if ME PD researchers attempted to isolate the effects of each of these variables within their studies, or, working across studies, examine the same or similar sets of components or characteristics. This argument is similar to Desimone's (2009) call for PD researchers to examine the same set of core characteristics or features across programs in order to better compare and understand the effects of each (i.e., content focus, active learning, coherence, duration, and collective participation). We believe ME PD researchers should do the same, with the addition of those features identified here as potentially also influential (i.e., facilitator background, tailoring to specific student population, and incorporation of critical friendships, self-evaluation, immersion experiences, and any other aspects of ME PD programs that emerge as potentially powerful). We do not mean to suggest these are the only features that should be studied, only that if more researchers attended to this same minimal set of features, then we might begin to understand whether, and if so how, each contributes to stronger ME PD.

Last but not least, researchers and PD designers (who are often one and the same) should be critically self-reflective as they articulate their underlying theories of both teacher learning and multicultural education. They should consider whether their interpretation of multicultural education is likely to assist teachers in interrogating simplistic representations of cultures and the broader sociopolitical contexts within which those cultures are situated. Formulations of

multicultural education that fail to do so not only hinder teachers from developing the critical lens necessary for pursuing educational justice, they also may support prejudicial thinking and behaviors (Schmeichel, 2012). Researchers should also consider how their particular theory of teacher learning might shape their interpretations of findings.

Conclusion

The studies reviewed here lack sufficient consistency across theoretical approaches, PD designs, and data collection methods to draw definitive conclusions about the characteristics of effective ME PD. However, taken together these studies point to a number of questions that ME PD researchers should investigate and PD providers may want to consider. First, how can ME PD both challenge teachers to reflect on inequities within education while also recognizing that some teachers may meet such discussions with defensiveness, reluctance to change or skepticism about the importance of ME? Second, how can providers strike a balance between providing specific knowledge about students' cultures, for instance through partnering with community members, and guarding against promoting stereotypes or broad generalizations?

ME PD researchers and providers should attend to both their underlying theories of multicultural education as well as theories of how teachers learn. These studies caution against assuming that raising awareness of diversity and inequities will naturally lead to transformed teaching practices or that teachers will develop culturally responsive lessons without specific guidance on how to connect cultural assets to their curriculum (e.g., Brown & Crippen, 2016; Lee et al., 2007). Theories of how teachers learn should also attend to contexts that support professional growth. For instance, does embedding PD programs within larger, district-wide initiatives related to ME augment teacher development? Finally, much greater attention should be paid to the range of starting points of educators and the likelihood that resistance or

disengagement may result from a number of challenges inherent in this type of PD. Some teachers may take inequities as a given condition (Lee et al., 2007) or feel they are already sufficiently addressing issues of diversity (Zozakiewicz & Rodriguez, 2007). Understanding how teachers reconsider their views of diversity and equity is a task for more than ME PD research alone. Developing this understanding would benefit everyone in the field of education seeking a future in which students from all backgrounds receive inclusive, equitable, and high-quality education.

References

* indicates studies included in the systematic literature review

Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. *Review of Educational Research*, *86*(1), 163-206.

*Aujla-Bhullar, S. (2011). Deconstructing diversity: Professional development for elementary teachers. *Diaspora, Indigenous, and Minority Education, 5*, 266-276.

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., & Shepard, L.A. (2010). Problems with the use of student test scores to evaluate teachers. EPI Briefing Paper No. 278. *Economic Policy Institute*.

Banks. (2004). Multicultural education: Historical development, dimensions, and practice. In J. A. Banks & C. A. Banks (Eds.), *Handbook of research on multicultural education* (2nd ed.) (pp. 3-29). San Francisco, CA: Jossey-Bass.

Banks, J. A., & Banks, C. A. (2004). *Handbook of research on multicultural education* (2nd ed.). San Francisco, CA: Jossey-Bass.

*Bishop, R., Berryman, M., Cavanagh, T., & Teddy, L. (2009). Te kotahitanga: Addressing educational disparities facing Māori students in New Zealand. *Teaching and Teacher Education, 25*(5), 734-742.

Blank, R. K., & de las Alas, N. (2009, June). Effects of teacher professional development on gains in student achievement: How meta-analysis provides scientific evidence useful to educational leaders. Washington, DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Documents/2009/Effects_of_Teacher_Professional_2009.pdf

*Brown, J. C., & Crippen, K. J. (2016). Designing for culturally responsive science education through professional development. *International Journal of Science Education, 38*(3), 470-492.

Cabrera, N. L., Milem, J. F., Jaquette, O., & Marx, R. W. (2014). Missing the (student achievement) forest for all the (political) trees: Empiricism and the Mexican American studies controversy in Tucson. *American Educational Research Journal, 51*(6), 1084-1118.

*Charity Hudley, A. H., & Mallinson, C. (2017). "It's worth our time": A model of culturally and linguistically supportive professional development for K-12 STEM educators. *Cultural Studies of Science Education, 12*(3), 637-660.

Cloonan, A., Fox, B., Ohi, S., & Halse, C. (2017). An analysis of the use of autobiographical narrative for teachers' intercultural learning. *Teaching Education, 28*(2), 131-144. doi:10.1080/10476210.2016.1212005

*Colombo, M. W. (2007). Developing cultural competence: Mainstream teachers and

    professional development. *Multicultural Perspectives, 9*(2), 10-16.

Cooper, H. (2017). *Research synthesis and meta-analysis: A step-by-step approach* (5th ed.).

    Thousand Oaks, CA: Sage Publications.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, S. (2009).

    Professional learning in the learning profession: A status report on teacher development

    in the United States and abroad. *Washington, DC: National Staff Development Council*.

*de la Garza, K. (2016). Pedagogical mentorship as an in-service training resource: Perspectives

    from teachers in Guatemalan rural and indigenous schools. *Global Education Review*,

    3(1), 45-65.

Dee, T. S., & Penner, E. K. (2017). The causal effects of cultural relevance: Evidence from an

    ethnic studies curriculum. *American Educational Research Journal*, *54*(1), 127-166.

*DeJaeghere, J. G., & Cao, Y. (2009). Developing US teachers' intercultural competence: Does

    professional development matter? *International Journal of Intercultural Relations*, *33*(5),

    437-447.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development:

    Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199.

Desimone, L., Smith, T. M., & Phillips, K. J. (2007). Does policy influence mathematics and

    science teachers' participation in professional development?. *Teachers College

    Record*, *109*(5), 1086-1122.

*Devereaux, T. H., Prater, M. A., Jackson, A., Heath, M. A., & Carter, N. J. (2010). Special

    education faculty perceptions of participating in a culturally responsive professional

    development program. *Teacher Education and Special Education, 33*(4), 263-278.

*Dimitriadou, C., Nari, E., & Palaiologou, N. (2012). E-learning teacher training courses for

     differentiated instruction in multicultural classrooms: Reflections upon the participants'

     experiences. *Journal of Educational Technology*, *9*(3), 14-26.

Echevarria, J., Vogt, M. E. & Short, D. J. (2016). *Making content comprehensible for English

     language learners: The SIOP model (5th ed.).* Boston, MA: Pearson.

Elmi, A. A. (2010). Peace education in Canada: Teacher perceptions of the cultivating peace

     education program. *Journal of Contemporary Issues in Education*, *4*(2), 41-54.

Gándara, P., Maxwell-Jolly, J., & Driscoll, A. (2005). Listening to teachers of English language

     learners: A survey of California teachers' challenges, experiences, and professional

     development needs. *Policy Analysis for California Education, PACE (NJ1)*.

García, S. S., & García, C. F. (2016). Transformative professional development and the

     promotion of literacy through culturally responsive pedagogy. *CATESOL Journal, 28*(1),

     175-194.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes

     professional development effective? Results from a national sample of teachers.

     American Educational Research Journal, 38(4), 915–945.

Garrett, H. J., & Segall, A. (2013). (Re) considerations of ignorance and resistance in teacher

     education. *Journal of Teacher Education*, *64*(4), 294-304.

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education*,

     53(2), 106-116.

Gooden, M. A., & O'Doherty, A. (2015). Do you see what I see? Fostering aspiring leaders'

     racial awareness. *Urban Education, 50*(2), 225-255.

*Grimberg, B. I., & Gummer, E. (2013). Teaching science from cultural points of intersection.

    Journal of Research in Science Teaching, *50*(1), 12-32.

Gulamhussein, A. (2013). Teaching the teachers: Effective professional development in an era of

    high stakes accountability. *Center for Public Education*, 1-47.

Harber, C., & Sakade, N. (2009). Schooling for violence and peace: How does peace education

    differ from 'normal' schooling? *Journal of Peace Education*, *6*(2), 171-187.

Hollins, E., & Guzman, M. T. (2005). Research on preparing teachers for diverse populations. In

    M. Cochran-Smith & K. M. Zeichner (Eds.), *Studying teacher education: The report of*

    the AERA panel on research and teacher education (pp. 477-548). Mahwah, NJ:

    Lawrence Erlbaum Associates, Inc.

*Hulan, N. F. (2015). Molding a culturally responsive literacy practice: Professional

    development within diverse schools. *Teacher Education and Practice, 28*(1), 60-74.

Jiang, B., & DeVillar, R. A. (2011). Effects of school and cultural contexts on student teaching

    abroad outcomes: Insights from U.S. student teachers in Belize, China, and Mexico.

    Action in Teacher Education, 33, 47-62.

*Johnson, C. C. (2011). The road to culturally relevant science: Exploring how teachers navigate

    change in pedagogy. *Journal of Research in Science Teaching, 48*(2), 170-198.

*Johnson, C. C., & Fargo, J. D. (2014). A study of the impact of transformative professional

    development on Hispanic student performance on state mandated assessments of science

    in elementary school. *Journal of Science Teacher Education, 25*(7), 845-859.

*Johnson, C. C., & Marx, S. (2009). Transformative professional development: A model for

    urban science education reform. *Journal of Science Teacher Education, 20*(2), 113-134.

*Katz, L., Inan, H. Z., Tyson, C., Dixson, A., & Kang, H. Y. (2010). Professional development

     for the early learning content social studies standards. *International Electronic Journal of*

     Elementary Education, 2(2), 261-285.

Kennedy, A. (2005). Models of continuing professional development: A framework for analysis.

     Journal of In-service Education, *31*(2), 235-250.

Kennedy, M. (2016). How does professional development improve teaching? *Review of*

     Educational Research, *86*(4), 945-980.

Ladson-Billings, G. J. (1995). Toward a theory of culturally relevant pedagogy. *American*

     Education Research Journal, 35, 465-491.

*Lee, O., Luykx, A., Buxton, C., & Shaver, A. (2007). The challenge of altering elementary

     school teachers' beliefs and practices regarding linguistic and cultural diversity in science

     instruction. *Journal of Research in Science Teaching, 44*(9), 1269-1291.

Lee, E., Menkart, D., & Okazawa-Rey, M. (2011). *Beyond heroes and holidays: A practical*

     guide to K-12 anti-racist, multicultural education and staff development. Washington,

     DC: Teaching for Change.

Little, J. W. (1987). Teachers as colleagues. In V. Richardson-Koehler (Ed.), *Educators'*

     handbook: A research perspective (pp. 491-518). New York, NY: Longman.

*Marshall, P. L., DeCuir-Gunby, J. T., & McCulloch, A. W. (2012). In search of professional

     dispositions that yield cultural relevance in primary grade pedagogy: A cautionary tale of

     one kindergarten teacher. *Journal of Praxis in Multicultural Education*, 7(1), Article 3.

*Marshall, P., & McCulloch, A. (2010). Un-bordering early mathematics pedagogy: Culture,

     content, and identity in critical professional development. *Boğaziçi University Journal of*

     Education, *27*(1), 1-21.

Marx, H., & Moss, D. (2011). Please mind the culture gap: Intercultural development during a

 teacher education study abroad program. *Journal of Teacher Education, 62*(1), 35-47.

Mayfield, V. M., & Garrison-Wade, D. (2015). Culturally responsive practices as whole school

 reform. *Journal of Instructional Pedagogies, 16*, 1-17.

McAllister, G., & Irvine, J. J. (2000). Cross cultural competency and multicultural teacher

 education. *Review of Educational Research, 70*(1), 3-24.

*McCulloch, A. W., & Marshall, P. L. (2011). K–2 teachers' attempts to connect out-of-school

 experiences to in-school mathematics learning. *Journal of Urban Mathematics

 Education, 4*(2), 44-66.

*Messiou, K., & Ainscow, M. (2015). Responding to learner diversity: Student views as a

 catalyst for powerful teacher development? *Teaching and Teacher Education, 51*, 246-

 255.

Mette, I. M., Nieuwenhuizen, L., & Hvidston, D. J. (2016). Teachers' perceptions of culturally

 responsive pedagogy and the impact on leadership preparation: Lessons for future reform

 efforts. *International Journal of Educational Leadership Preparation*, *11*(1).

Miles, J., & Bailey-McKenna, M. C. (2016). Giving refugee students a strong head start: The

 LEAD program. *TESL Canada Journal*, *33*, 109-128.

Miles, K. H., Odden, A., Fermanich, M., & Archibald, S. (2004). Inside the black box of school

 district spending on professional development: Lessons from five urban districts. *Journal

 of Education Finance, 30*(1), 1–26.

*Molle, D. (2013). Facilitating professional development for teachers of English language

 learners. *Teaching and Teacher Education*, *29*, 197-207.

National Center for Education Statistics. (2018, May). Public high school graduation rates.

> *National Center for Education Statistics*. Retrieved from

> https://nces.ed.gov/programs/coe/indicator_coi.asp

Noblit, G. W. (2018). Meta-ethnography: Adaptation and return. In L. Urrieta Jr. & G. W. Noblit

> (Eds.), *Cultural constructions of identity: Meta-ethnography and theory* (pp. 34-50).

> Oxford University Press.

Noblit, G. W., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies* (Vol.

> 11). Newbury Park, CA: Sage.

*Orosco, M. J., & Abdulrahim, N. A. (2017). Culturally responsive professional development for

> one special education teacher of Latino English language learners with mathematics

> learning disabilities. *Insights into Learning Disabilities, 14*(1), 73-95.

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and

> practice. *Educational Researcher*, *41*(3), 93-97.

Paris, D., & Alim, H. S. (Eds.). (2017). *Culturally sustaining pedagogies: Teaching and learning

> for justice in a changing world*. New York, NY: Teachers College Press.

Parkhouse, H. & Rong, X. L. (2016). Inequalities in China's compulsory education: Progress,

> inadequacies, and recommendations. In S. Guo & Y. Guo (Eds.), *Spotlight on China:

> Changes in education under China's market economy.* Rotterdam, Netherlands: Sense.

Parkhouse, H., Turner, A. M., Konle, S., & Rong, X. L. (2016). Self-authoring the meaning of

> student teaching in China: Impacts on first-year teaching practices. *Frontiers: The

> Interdisciplinary Journal of Study Abroad*, *28*, 78-98.

*Papanastasiou, E. C., & Conway, P. F. (2002). Teacher professional development through LATTICE an international-intercultural project. *Studies in Educational Evaluation, 28*(4), 315-328.

Petticrew, M., & Roberts, H. (2008). Systematic reviews-do they 'work' in informing decision-making around health inequalities? *Health Economics, Policy and Law*, *3*(2), 197-211. doi:10.1017/S1744133108004453

*Powell, R., Cantrell, S. C., & Correll, P. (2016). Operationalizing culturally responsive instruction: Preliminary findings of CRIOP research. *Teachers College Record, 118*(1), 1-46.

*Prater, M. A., Wilder, L. K., & Dyches, T. T. (2008). Shaping one traditional special educator preparation program toward more cultural competence. *Teaching Education, 19*(2), 137–151. doi:10.1080/10476210802040765

*Psalti, A. (2007). Training Greek teachers in cultural awareness: A pilot teacher-training programme—implications for the practice of school psychology. *School Psychology International, 28*(2), 148-162. doi:10.1177/0143034307078090

Quezada, R. L. (2004). Beyond educational tourism: Lessons learned while student teaching abroad. *International Education Journal*, *5*(4), 458-465.

*Rubel, L. H., & Chu, H. (2012). Reinscribing urban: Teaching high school mathematics in low income, urban communities of color. *Journal of Mathematics Teacher Education, 15*(1), 39-52. doi:10.1007/s10857-011-9200-1

Saint-Hilaire, L. A. (2014). So, how do I teach them? Understanding multicultural education and culturally relevant pedagogy. *Reflective Practice, 15*(5), 592-602.

*Savage, C., Hindle, R., Meyer, L. H., Hynds, A., Penetito, W., & Sleeter, C. E. (2011).

    Culturally responsive pedagogies in the classroom: Indigenous student experiences across

    the curriculum. *Asia-Pacific Journal of Teacher Education, 39*(3), 183-198.

    doi:10.1080/1359866X.2011.588311

Schmeichel, M. (2012). Good teaching? An examination of culturally relevant pedagogy as an

    equity practice. *Journal of Curriculum Studies, 44*(2), 211-231.

    doi:10.1080/00220272.2011.591434

*Schniedewind, N. (2001). Embracing the rainbow: An integrated approach to diversity

    education. *Multicultural Perspectives, 3*(1), 23–27. doi:10.1207/S15327892MCP0301_5

*Schroeder, J. L., Plata, M., Fullwood, H., Price, M., & Dyer Sennette, J. (2013). Increasing the

    cultural competence of professionals via online training. *National Forum of Multicultural*

    Issues Journal, 10(1), 1-13.

Shaha, S., Glassett, K., Copas, A., & Huddleston, T. L. (2016). Maximizing educator

    enhancement: Aligned seminar and online professional development. *Contemporary*

    Issues in Education Research, 9(1), 15-22.

*Sleeter, C. E. (Ed.). (2011). *Professional development for culturally responsive and*

    relationship-based pedagogy. Black studies and critical thinking. Volume 24. New York,

    NY: Peter Lang.

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is

    real: Anthropological and historical perspectives on the social construction of

    race. *American Psychologist, 60*(1), 16-26. doi:10.1037/0003-066X.60.1.16

*Smith, N. L., & Bahr, M. W. (2014). Increasing cultural competence through needs assessment and professional development. *Professional Development in Education, 40*(1), 164-181. doi:10.1080/19415257.2013.776618

Smith, E. J., & Harper, S. R. (2015). Disproportionate impact of K-12 school suspension and expulsion on Black students in southern states. *Center for the Study of Race and Equity in Education.* Philadelphia, PA: University of Pennsylvania.

Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, *62*(4), 339-355.

*Timmons-Brown, S., & Warner, C. (2016). Using a conference workshop setting to engage mathematics teachers in culturally relevant pedagogy. *Journal of Urban Mathematics Education, 9*(1), 19-47.

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). Teacher professional learning and development: Best evidence synthesis iteration (BES). Auckland, New Zealand: University of Auckland. Retrieved from http://www.educationcounts.govt.nz/__data/assets/pdf_file/0017/16901/TPLandDBESentireWeb.pdf

*Trifiro, A. J. (2017). Transforming teachers' practice through professional development: Culturally sustaining pedagogical changes in support of English language learners. In *Culturally Sustaining and Revitalizing Pedagogies: Language, Culture, and Power* (pp. 269-287). Emerald Publishing Limited.

Trilokekar, R. D., & Kukar, P. (2011). Disorienting experiences during study abroad: Reflections of pre-service teacher candidates. *Teaching and Teacher Education*, *27*, 1141-1150. doi:10.1016/j.tate.2011.06.002

*Tucker, C. M., Porter, T., Reinke, W. M., Herman, K. C., Ivery, P. D., Mack, C. E., & Jackson, E. S. (2005). Promoting teacher efficacy for working with culturally diverse students. *Preventing School Failure: Alternative Education for Children and Youth, 50*(1), 29-34.

Urrieta Jr, L., & Noblit, G. W. (Eds.). (2018). *Cultural constructions of identity: Meta-ethnography and theory*. Oxford, UK: Oxford University Press.

Valenzuela, A. (1999). *Subtractive schooling: U.S.-Mexican youth and the politics of caring*. Albany, NY: State University of New York Press.

*Voltz, D. L., Brazil, N., & Scott, R. (2003). Professional development for culturally responsive instruction: A promising practice for addressing the disproportionate representation of students of color in special education. *Teacher Education and Special Education, 26*(1), 63-73. doi:10.1177/088840640302600107

*Watkins, M., & Noble, G. (2016). Thinking beyond recognition: Multiculturalism, cultural intelligence, and the professional capacities of teachers. *Review of Education, Pedagogy, and Cultural Studies, 38*(1), 42-57. doi:10.1080/10714413.2016.1119642

Wing Sue, D., Torino, G. C., Capodilupo, C. M., Rivera, D. P., & Lin, A. I. (2009). How White faculty perceive and react to difficult dialogues on race: Implications for education and training. *The Counseling Psychologist*, *37*(8), 1090-1115. doi:10.1177/0011000009340443

*Wiseman, A., & Fox, R. K. (2012). Supporting teachers' development of cultural competence

      through teacher research. *Action in Teacher Education, 32*(4), 26-37.

      doi:10.1080/01626620.2011.627033

Zepeda, S. J. (2012). *Professional development: What works* (2nd ed.). Larchmont, NY: Eye on

      Education.

Zirkel, S. (2008). The influence of multicultural educational practices on student outcomes and

      intergroup relations. *Teachers College Record*, *110*(6), 1147-1181.

*Zozakiewicz, C., & Rodriguez, A. J. (2007). Using sociotransformative constructivism to create

      multicultural and gender-inclusive classrooms: An intervention project for teacher

      professional development. *Educational Policy, 21*(2), 397-425.

      doi:10.1177/0895904806290126

Appendix

Summary of Studies

Study	School/district student demographics	Specific Student Population	Description of Program	Major Contributions of Study to the Knowledge Base on ME PD
Aujla-Bhullar (2011)	1 school in Calgary, Canada with a majority of students from South Asian backgrounds	Muslim and Sikh students	1-day workshop including identity reflections in small groups, guest speakers, and field trip to a Mosque and a Sikh temple	Participants reported growth, but many also demonstrated discomfort and a fear of offending; some pushed away thoughts and feelings to maintain a positive environment
Bishop et al. (2009); Savage et al. (2011); Sleeter (2011)	12 New Zealand schools where more than 20% of students identified as Māori	Māori students	Te Kotahitanga: An initial induction workshop, classroom observations and feedback, problem-solving sessions using student data, & coaching sessions. Māori students & parents primarily drove the PD content	Māori students collaborated on the PD framework. Teachers reported transformations after hearing from students. After the PD, students felt teachers valued and understood their culture. Literacy, numeracy, attendance, and engagement increased following the PD.
Brown & Crippen (2016)	Large, culturally and linguistically diverse school district in the Southeastern US	CLD students	6 Saturday sessions that began with reflecting on the student population, followed by analyzing exemplary lessons, and finally creating an action plan aligned with research	Teachers recognized inequities in their own classroom (e.g., gendered participation) and envisioned science-specific culturally responsive lessons; some limited impacts due to lack of teacher knowledge of students' specific cultures
Charity Hudley & Mallinson (2017)	K-12 STEM educators in Maryland and Virginia	Linguistically diverse learners	Workshops focused on language, pedagogical practices, and student assessment related to linguistic diversity and language variation, particularly for Afr. Am. students	Teachers learned how to connect culture to STEM, stopped correcting non-standard varieties of English without explanation, and became more mindful of science jargon
Colombo (2007)	1 school district in which 16% of students spoke Spanish as a first language and 20% were considered CLD	Spanish-speaking students	Parent coordinators offered after-school courses to teachers on cultural awareness, diversity and schooling, ESL teaching practices, second language literacy & instructional strategies. Teachers, in pairs, attended family literacy nights as well.	A sense of being lost helped teachers understand students; the resistance of two paired teachers persisted as they were able to reinforce one another's deficit perspectives toward parents and skepticism toward ME
de la Garza (2016)	Rural and indigenous school communities in Baja Verapaz and Quiche in Guatemala	Indigenous language speaking students	Teachers received mentorship to help with their bilingual and intercultural education instruction skills and to help change ideas of cultural inferiority and superiority.	Mentorships helped teachers gain ideas for teaching L1 (indigenous languages), classroom materials, games, and songs. They also gained a better understanding of the National Base Curriculum of Guatemala.

DeJaeghere & Cao (2009)	7 urban schools in the US; 38% students of color at the start of study and 50% at the end	No	Tailored for each school according to the staff's aggregated scores on Intercultural Development Inventory	Growth in intercultural competence (ICC), especially understanding of cultural similarities & differences, can be obtained through tailoring PD to particular ICC levels of each school.
Devereaux et al. (2010); Prater et al. (2008)	Special education faculty members at a Western U.S. university with 9% teacher candidates identified as ethnically diverse & about 50% fluent in a 2nd language	Teacher candidates	Faculty members were taught CREDE (2009) standards and evaluation methods (i.e. language development, contextualization, and instructional conversations). They also conducted home visits to teacher candidates of ethnically diverse/bilingual backgrounds.	Faculty enjoyed the home visits to pre-service teachers' homes more than the lecture portions and other activities; some "disliked the lack of honest self-reflection by fellow faculty members."
Dimitriadou, Nari, & Palaiologou (2012)	Developed by faculty at the Aristotle University of Thessaloniki and included 16 teachers from Western Macedonia teaching in schools with multicultural classes	Students of diverse socio-cultural backgrounds	A 6-week distance e-learning course on differentiated instruction. The goal of the thematic unit studies was to help teachers use a poster as a visual narrative inquiry text in order to enable students to construct personally meaningful knowledge	Distance learning can help teachers feel more confident in their ability to collect appropriate educational materials, communicate with their students, and implement student-centered methods. All teachers expressed an interest in future training programs and staying up-to-date with teaching techniques.
Grimberg & Gummer (2012)	The northwest and southeast regions of Montana, which include the Flathead, Crow, and Northern Cheyenne Reservations.	American Indian students	A 5-year long PD program comprised of face-to-face workshops and online interactions. Tribal leaders drove PD content with the aim to develop teachers' knowledge of tribal cultures and science content, as well as teaching methods to bridge the two.	Involving tribal leaders in PD development and facilitation helped teachers learn more about their students' cultures and connections to science. Teachers became more confident in implementing equitable instruction & increased time spent helping students make connections between home and science.
Hulan (2015)	4 elementary schools in which 70-96% of the students received free or reduced-price lunch	Diverse students	Teachers participated in a PD "book club." They read articles on topics related to culturally responsive instruction, completed tasks in their classrooms, kept journals about their reading, were videotaped, and watched their video recordings with researchers to discuss discourse practices.	Teachers demonstrated growth in knowledge of culturally responsive instructional strategies. However, some teachers struggled to create lessons based on students' background knowledge.
Johnson (2011); Johnson & Fargo (2014); Johnson & Marx (2009)	Urban-centered district in the Southwestern United States in which the majority of students were identified as ELL	ELLs	Transformative Professional Development (TPD): Two-year program comprised of 2-week long summer workshops, graduate courses, monthly grade-level support sessions, a conversational Spanish course, and home visits	Participants transformed their practice to focus on culturally relevant science pedagogy resulting in a more effective instructional environment. Teachers used more cooperative learning, scaffolding, inquiry, and built relationships with students. Participation in TPD had a significant impact

				on student achievement with the percentage of proficient students growing from 25% at baseline to 67% at the end.
Katz, Inan, Tyson, Dixson, & Kang (2010)	Early childhood educators across Ohio	Diverse early childhood students	Modules focusing on the social studies standards and effective pedagogy for young children presented in three sessions.	Educators reported using more multicultural literature to support teaching social studies at the early childhood level.
Lee, Luykx, Buxton, & Shaver (2007)	6 elementary schools in urban district in Southeastern US with 70% FRL, 25% ELLs, 57% Hispanic, 30% Black, 11% White non-Hispanic, and 2% Asian or Native American.	ELLs	Four full-day workshops on the following topics: inquiry-base science, English literacy, home language and culture incorporation, and lesson workshopping in cross-grade small groups. In addition, a speaker from Haiti talked about her experiences as an immigrant student in U.S. schools	Deficit perspectives of immigrant parents didn't always change; Teachers thought incorporation of home language & culture were important but lacked adequate knowledge to do so; no change in how often teachers used home language/culture in instruction
Marshall et al. (2012); Marshall & McCulloch (2010); McCulloch & Marshall (2011)	Elementary schools in a large urban school district in a Southeastern state of the US.	Students of color	Nurturing Mathematics Dreamkeepers program (NMD): Four 2-day sessions throughout the school year and a one week intensive summer institute covering the following: mathematics content knowledge, critical awareness of how children understand key early elementary mathematics concepts, the influence of cultural factors on the teaching-learning process, and analysis of student work samples and classroom videos. Participants were required to observe each other teaching two consecutive mathematics lessons.	Teachers' actions fell into 4 categories: clarifying mathematical context, making mathematics relevant, introduction of formal mathematics, and validating students' mathematical contributions. However, they also found: a low rate of teachers connecting to students' out-of-school lives, low levels of mathematics content and pedagogical content knowledge, colorblind and deficit perspectives, and the belief that mathematics is culture-free.
Messiou & Ainscow (2015)	8 diverse schools in the cities of Hull, England, Manchester, England, Lisbon, Portugal, and Madrid, Spain	Diverse students	Each school had a team of 3 teachers that used the process of lesson study to explore how to make their classrooms more inclusive. They also incorporated students' opinions and suggestions.	Teachers changed some of their instructional practices as a result of hearing from their students. They also went from "adopting" and "creating" categories of students to "rethinking" those categories as a result of hearing from students.
Mette, Nieuwenhuizen, & Hvidston (2016)	1 high school in a Midwestern city in which 71% of students were White and 29% were students of color	No	The multicultural committee researched the achievement gap, white privilege, and CRP and then led scripted 45-minute sessions with their PLCs consisting of norm-setting, content delivery, discussion, and reflection	Increases in both ELA and math scores of students of color, but decrease in science. Teachers of electives and special education viewed the PD as having a greater impact on instruction than core content teachers did.
Molle (2013)	5 PD teams made up of 21 teachers and administrators from	ELLs	Case study of 1 facilitator in the Content and Language Integration as a Means of Bridging Success (CLIMBS)	The facilitator fostered a productive learning environment by navigating tensions in ways that preserved collegiality among

	different districts in a Midwestern U.S. state		program, which consisted of 5 monthly, day-long meetings	participants, redirected negative discourses about language minority students, & made space for political discussions regarding ELLs
Orosco & Abdulrahim (2017)	Southwestern U.S. urban school of 453 students: 55% Latinx, 22% African American, 14% White, 5% Asian, and 4% other. 75% FRL	ELLs & students with math learning disabilities	Case study of one elementary special education teacher of Latinx ELLs. She identified her instructional needs and then participated in a PD tailored to those needs. After this training, the teacher received monthly booster workshops.	The teacher was able to provide instruction that built on students' background knowledge and incorporated explicit language instruction.
Papanastasiou & Conway (2002)	Teachers and administrators from three school districts near Michigan State	Culturally diverse students	LATTICE, a long-term international intercultural PD program. Workshops included a topic presentation, small group discussion, an international food break, whole group discussion, and a survey at the end	Many teachers' thoughts on international issues changed over time, which influenced their teaching practices.
Powell, Cantrell, Malo-Juvera, & Correll (2016)	Teachers from 4 schools in/near a mid-sized Midwestern city; 456 student participants: 87.3% FRL & 28% ELLs	Historically underrepresented student populations	3 workshops were held before the school year on elements of the Culturally Responsive Instruction Observation Protocol (CRIOP). During the school year, teachers received individual classroom coaching, onsite PD, & instructional planning support.	Teachers grew significantly in CRIOP implementation. Students with teachers identified as *high implementers* scored significantly higher than those with *low implementer* teachers on reading/math tests.
Psalti (2007)	70 primary and secondary teachers at state schools in three prefectures of Northern Greece	CLD students	Workshop based on ASK model: Awareness, Skills and sensitivity, and Knowledge. There were three phases: needs assessment of all participants, teacher training in cultural awareness, and intervention in the classroom.	Teachers' ability to communicate with and effectively approach students improved; Teachers identified PD weaknesses as organizational difficulties and hard meeting hours.
Rubel & Chu (2012)	2 NYC high schools: one in which 2/3 of students identified as Latinx & 1/3 Black/African American. In the other, 90% of students identified as Black/African-American & 10% Latinx.	Students of color	Two 5-day summer institutes, two 1-day fall seminars, and 12-15 school-based teacher meetings in which teachers were introduced to Culturally Relevant Mathematics Pedagogy (CureMap), which is a framework that addresses access, achievement, identity and power.	Two of the 7 teachers' practices were highlighted as evidence of CureMap, while the remaining 5 teachers continued to use tasks of low-level cognitive demand; 4 of the 7 teachers nearly exclusively offered students only 2 ways to participate: listening and practicing.
Schniedewind (2005)	A semi-rural community in New York with a student body that is primarily White and about 15% FRL	Students of color	A voluntary 30-hour PD course in diversity education and then 2 recorded follow-up discussions with the educators	Teachers supported students of color by validating their racial identities and creating safe classrooms to talk about color, bias, and racism. Teachers intentionally encouraged White students to talk about race and challenge white privilege.
Schroeder, Plata, Fullwood, Price,	Not specified (Special education assessment	CLD students	7 modules were related to diversity issues in education and 2 were on strategies for assessing diverse learners	Online training programs are cost-efficient, highly individualized, and allow repeated

& Dyer Sennette (2013)	professionals participated in 9 online training modules)		through an online teaching platform. Each training module consisted of a lecture, an assignment, and a short-answer quiz.	reviewing of material. They can easily be updated with new material, and they provide anonymity for those who may find face-to-face training intimidating or insulting.
Smith & Bahr (2014)	School-based mental health center located in an urban, southern US school district: 86% African American, 12% White, and 2% other nationalities.	Latinx ELLs	PD for school psychologists and social workers. Began with needs assessment and grounded PD in personnel's perceptions of their strengths and weaknesses. The PD used the 3-stage model of awareness, knowledge, and skills; included a presentation on cultural characteristics of the Latinx population, second language acquisition, best practices in working with translators, and patterns of Latinx parental involvement in the school	Half of the participants said the PD increased their awareness of their own cultural background, attitudes, and biases; 21% indicated that the training was helpful in increasing knowledge of cultural group differences; 35% cited increases in cultural communication and skills; providers cited difficulties including all of the identified needs in the training
Timmons-Brown & Warner (2016)	Teachers from 4 elementary schools and 1 middle school with majority Black student populations and 1 middle school with majority White student population	Underserved student populations	A 2-day conference designed for elementary, middle, and high school math teachers. Teachers were invited to attend as school teams and were provided tools and instruction for classroom use.	Reported increases in knowledge of culturally relevant terminology and use of CRP practices. Participants began to see students as individuals. Networking with others gave participants new language to communicate experience with colleagues.
Trifiro (2017)	Arizona, which was an English-only instruction state at the time; no details on student demographics	ELLs	PD for urban teachers aiming to improve teaching practices for ELLs. PD included 2 courses alongside 2 semester-long practicum experiences that integrated ongoing coaching and observational feedback. Teachers chose 2 ELL students to focus on as a case study. Teachers charted personal growth and kept online journals.	Teachers identified the importance of making connections to ELLs, found ways to integrate culture with content, used first language to support ELLs, and incorporated culturally sustaining pedagogy into their practice.
Tucker, Porter, Reinke, Herman, Ivery, Mack, & Jackson (2005)	Teachers from 6 schools in a Southeastern region of the US that had similar demographics & had received a "D" on a statewide assessment of student performance	Culturally diverse students	Teachers in the treatment conditions received a 6-hour workshop designed to inform teachers about effective methods and strategies for teaching culturally diverse students. These teachers were also invited to attend a follow-up session 9 weeks later.	Teacher participants scored higher on posttest items measuring general teaching and culturally sensitive self-efficacy. In addition, teachers reported an increase in their knowledge about culturally diverse students.
Voltz, Brazil, & Scott (2003)	20 general and 13 special education teachers from a large, metropolitan school district. The majority of teachers had a student	Special education & diverse students	Culturally Responsive Instruction for Special Population (CRISP) entailed a three-day seminar where teachers learned about James Bank's multicultural education framework and explored issues related to disability versus cultural difference. Participants then set goals	General education teachers said they felt better prepared to address the educational needs of culturally diverse students, collaborate with parents from diverse cultures, distinguish culturally-based learning

	population where 50% or less were minority students.		and then planned and implemented PD activities to accomplish them. PD activities included reading groups, action research projects, and curriculum development activities.	differences from disabilities, teach with a multicultural perspective, and consider a wide range of factors when making a referral. Both general and special education teachers agreed that their participation increased their knowledge and skills with respect to teaching with a multicultural perspective.
Watkins & Noble (2016)	14 primary and secondary schools in Australia varying in socio-economic status, geographic location (urban and rural), and the number of students who speak a language other than English at home	Students with a language background other than English	Rethinking Multiculturalism/Reassessing Multicultural Education (RMRME) Project: 5 schools developed action research teams composed of one administrator, one ESL teacher, and other teachers. The teams read and discussed three texts and unpacked concepts such as identity, transnationalism, and globalization, in the first stage. In the second they conducted action research cycles.	Few schools conducted action research projects that did more than revamp their current multicultural days and field trips. One school critically reflected on their practices and changed their ways. Some school projects faced teachers who resisted change, were unwilling participate, and took an anti-intellectual stance on the professional discourse of teaching.
Wiseman & Fox (2012)	Not specified	Diverse students	Teachers participated in ten face-to-face classes and four online discussions that aimed to expanded their understanding of diversity and multicultural education and apply it to their own practice. These included four experiences: (a) defining and understanding ideas about culture; (b) developing and reflecting individual cultural autobiographies; (c) experiencing and considering culture through simulation; and (d) researching a culturally relevant topic in their educational settings.	Teachers used PD to explore their own understanding of culture, and action research helped inform pedagogy. Some teachers were resistant to ideas related to multiculturalism and diversity.
Zozakiewicz & Rodriguez (2007)	Southwest borderlands of the US. All schools had a majority Latinx student population	Culturally diverse girls from impoverished schools	A 2-week-long summer institute: The first week focused on the integration of science and learning technologies, and the second focused on the integration of science, math, engineering, and learning technologies. Teachers also participated in monthly meetings to discuss progress and annual, daylong workshops that covered teacher requested content.	Although many reported growth, some teachers felt that the work they were doing was already multicultural and student-centered and thus did not need to change. In addition, one veteran teacher was not willing to transform her practices at this point in her career.

Note. CLD = culturally and linguistically diverse; CREDE = Center for Research on Education, Diversity, and Excellence; CRP = culturally responsive/relevant pedagogy; ELA = English language arts; ELL(s) = English language learner(s); ESL = English as a second language; FRL = free and reduced-price lunch; L1 = first language; PLCs = professional learning communities; STEM = science, technology, engineering, and math. CREDE (2009) as cited in Devereaux et al. (2010)