Check for updates

Version of Record at: https://doi.org/10.1177/23328584241310429

AERA Open
January-December 2025, Vol. 11, No. 1, pp. 1–23
DOI: 10.1177/23328584241310429
Article reuse guidelines: sagepub.com/journals-permissions
© The Author(s) 2025. https://journals.sagepub.com/home/ero

A Holistic Examination of How Professional Learning and Curriculum Relate to Ambitious and Culturally Relevant Instruction and Student Engagement

Laura M. Desimone
University of Delaware

Nick Bell
University of Albany

Arielle Lentz
University of Delaware

Kirsten L. Hill
Kirsten Lee Hill Consulting

Latrice Marianno
Southern Illinois University-Edwardsville

To shape professional learning (PL) and curriculum interventions for successfully transforming instruction, we need to better understand how multiple aspects of the system work. Applying structural equation modeling with a sample of 437 teachers in 153 middle and high schools in 11 districts serving 50% or more Black or Latinx students, we found that teachers are significantly more likely to use ambitious and culturally responsive (CR) instruction in the classroom and that student engagement increases when PL and curriculum adoption are (1) targeted toward ambitious and CR instruction, (2) provide clear, specific directions to teachers, (3) are aligned with each other, (4) are supported by teachers, and (5) are accompanied by incentives. PL worked through fostering stronger beliefs in and confidence using CR instruction. The curriculum and PL had both independent and interactive effects on bolstering the use of ambitious and CR instruction.

Keywords: *curriculum, educational reform, professional development, instructional practices, educational policy, school/teacher effectiveness, structural equation modeling, survey research*

Introduction: Significance of Understanding How Curriculum and Professional Learning Influence Teachers and Students

Understanding how curriculum and professional learning shape instructional change and how that change influences students is fundamental to many educational improvement efforts. Research is clear that the effectiveness of professional learning (PL) increases when it is aligned with the required curriculum (Penuel et al., 2011) but provides less evidence on the extent to which successful implementation of a new curriculum depends on aligned high-quality PL. Relatedly, we do not have a clear sense of the features of the policy environment that need to be in place for curriculum and PL to shape instruction effectively. Additionally, there has been an increased focus on leveraging curriculum and PL to foster more ambitious instruction—which challenges students to analyze, synthesize, and apply knowledge rather than exclusively memorize and perform procedures—and culturally responsive (CR) teaching—which recognizes and draws on students' backgrounds, knowledge, and experiences to shape instruction. However, we know little about the mechanisms through which PL works to foster shifts in CR teaching, and more research is needed on the relationship between student engagement and ambitious, CR, and traditional instruction.

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

Version of Record at: https://doi.org/10.1177/23328584241310429

Scholars suggest that teacher beliefs in the value of CR instruction and their confidence in using CR approaches are likely instrumental in bolstering their use of it in the classroom (eg, Bray-Clark & Bates, 2003; Civitillo et al., 2019; Comstock et al., 2023). And reformers suggest that use of ambitious and CR instruction will foster greater student engagement, a critical precursor to improved learning (Markowitz, 2018) and a fundamental goal of the push to move away from the predominant reliance on procedural instruction (Ginsberg, 2015; Hill, 2009; Nykiel-Herbert, 2010).

Our study provides a unique opportunity to empirically examine this system of relationships, fundamental to current instructional reform yet rarely examined holistically. Using structural equation modeling, we are able to examine multiple components of the instructional system concurrently to answer a series of related research questions to better understand how curriculum and PL relate to instruction and student engagement: (1) How does curriculum and PL with specific high-quality features relate to teachers' use of ambitious, CR, and traditional instruction? (2) Are instructional change efforts more successful when a new curriculum is coupled with aligned PL? (3) Is the relationship between PL and instruction mediated by teachers' confidence and beliefs around CR instruction? (4) Is the use of ambitious and CR instruction related to greater student engagement?

In our study, we estimate a comprehensive set of relationships to test the following hypotheses: (1) curriculum and PL with certain desirable policy features relate independently to more use of ambitious and CR and less traditional instruction; when both are in place, they have even stronger relationships; (2) ambitious and CR instructional approaches relate to greater student engagement, whereas reliance on traditional instruction is associated with decreased engagement; (3) PL with desirable policy features is associated with stronger beliefs and confidence around CR instruction; and (4) these stronger beliefs and confidence relate to greater use of CR and ambitious instruction, which increase student engagement.

Our study contributes to the literature by modeling these complex relationships simultaneously using structural equation modeling. Further, ours is the only study that we know of that explicitly measures features of curriculum and PL in multiple districts across the country to model both direct and interactive relationships with ambitious and CR instruction—important because it comes closer to approximating the actual real-world complex system of relationships. Further, we know of no other study that provides larger-scale evidence to document whether PL on CR instruction is related to teachers' beliefs and confidence in using CR approaches and whether CR instruction relates to student engagement. Our work also provides evidence for the underdocumented hypothesis that CR instruction works better in the context of ambitious instruction. And we ground our analysis in the policy attributes theory (Porter et al., 1986), which allows us to assess the relationship of instruction with PL and curriculum with a set of key policy features of specificity, authority (teacher buy-in), consistency, power (rewards and sanctions), and stability. We examine these questions systematically through a set of structural equation models with a sample of 437 mostly middle school mathematics teachers in 11 districts across the country, each serving 50% or more Black and Latinx students.

Drawing on the Literature: Links Between Policy Features of PL and the Curriculum and Teacher Beliefs, Confidence, Instruction, and Student Engagement

Effects of Curriculum and PL on Instruction

Policymakers and reforms have relied heavily on both curriculum and PL as mechanisms for instructional improvement. Curriculum plays a critical role in shaping what teachers cover in the classroom and the instructional approaches they use (eg, Agodini & Harris, 2010; Polikoff, 2021; Remillard et al., 2014). The design of curricular materials shapes practice by emphasizing specific content, offering ideas for particular instructional practices, and designating the role of teachers in instruction (Agodini & Harris, 2010; Ball & Cohen, 1996; Remillard et al., 2014). Alignment of the curriculum with other elements of the policy system, such as content standards, is essential for providing teachers with a clear path to implementation (Pak et al., 2020; Polikoff, 2015).

Many scholars argue that if curriculum adoption is not accompanied by guidance and support for teachers, it will be insufficient to ensure changes in instruction (eg, Alozie et al., 2010; Polikoff, 2021). Curriculum enactment can vary for a number of reasons. Curricular materials may be misinterpreted, especially if there is little guidance embedded within the curriculum (Alozie et al., 2010). Teachers need support in understanding how the curriculum relates to content standards and how to adapt it to meet the needs of their particular students (eg, Ball & Cohen, 1996; Freeman & Porter, 1989; Penuel et al., 2011; Yang et al., 2020).

PL is instrumental in addressing these needs and providing extra support to help teachers make the changes required by new curricula. Teacher PL has been the subject of hundreds of studies, including rigorous randomized control trials (e.g., Garet et al., 2008; 2016) ; and while results vary, there is substantial evidence that high-quality PL can change what and how teachers teach (Agodini & Harris, 2010; Desimone, 2009; Garet et al., 2001; Kennedy, 2016). These effects have been well documented in elementary, middle, and high school in studies across both science, technology, engineering, and math (Greenleaf et al., 2011; Lynch et al., 2019; Osborne et al., 2019; Penuel et al., 2011; Roschelle et al., 2010) and English language arts (ELA) (Babinski et al., 2018; Masters et al., 2010).

Linking PL directly to curriculum has the potential to make both the PL and curriculum more effective. For

Version of Record at: https://doi.org/10.1177/23328584241310429

example, Taylor et al. (2015) used a clustered randomized, controlled design to show that combining research-based science curricular materials with curricular-aligned PL had a positive impact on improving teacher practice, which, in turn, improved student achievement for high school students. Other scholars used meta-analyses to show that curricular-aligned PL has stronger effects than PL or curriculum used alone (eg, Lynch et al., 2019; Pellegrini et al., 2021). Similarly, a review of randomized, controlled trials of PL interventions concluded that activities closely aligned with curricular lessons were the ones most likely to result in consistent and longer-term implementation (Desimone & Garet, 2015).

Providing PL focused explicitly on the curriculum can offer opportunities for teachers to better understand instructional strategies prioritized in the curriculum as well as curricular goals (eg, Alozie et al., 2010). Additionally, curriculum-aligned PL may enable teachers to better situate their learning of the curriculum within their school or district context (eg, Ball & Cohen, 1996; Davis & Krajcik, 2005). Each of these benefits can be instrumental in addressing misconceptions and in supporting teachers to enact the curriculum as intended.

This literature forms the basis for our first two research questions: How does curriculum and PL with specific high-quality features relate to teachers' use of ambitious, CR, and traditional instruction? and Are efforts more successful when a new curriculum is coupled with aligned PL?

Which Characteristics of PL and Curriculum Matter? Applying the Policy Attributes Theory

Our work builds on previous studies of curriculum and PL by examining how a set of key *features* of curriculum and PL matter. We use the policy attributes theory (Porter, 1994) as a framework for identifying the curriculum and PL attributes critical for fostering successful implementation. The framework posits five dimensions through which policies, in this case PL, gain influence over practice—specificity, authority, consistency, power, and stability. We adopt the view that it is the respondents' perceptions of the attributes—how they individually experience them—that matter for determining behavior rather than the policy as written (see Desimone, 2002).

The specificity, or level of detail of the guidance in the curriculum and the PL, determines how much teacher invention is required (Ball & Cohen, 1996); for example, sample lesson plans can provide a concrete understanding of how to enact the curriculum that reduces unwanted variation (eg, Penuel et al., 2011). Normative (buy-in) and institutional (leadership and resources) authority encourage teachers to take up the practices that are being promoted (eg, Polikoff, 2021; Remillard & Bryans, 2004). Consistency or coherence in the policy environment is also important (Hodge & Stosich, 2022; Spillane & Jennings, 1997; Spillane et al., 2022), and alignment of the curriculum with other initiatives and with PL ensures coherent guidance that facilitates implementation (e.g., Davis & Krajcik, 2005; Polikoff, 2021). Although results are mixed, research has shown that power in the form of incentives or sanctions can play a role in moving teachers to use a curriculum or take up what they learn in PL (eg, Dee & Wyckoff, 2015; L. Desimone, 2002; Yuan et al., 2013). Finally, the stability of the curriculum is crucial in supporting efforts to shift teachers' practice (eg, Berends et al., 2002). Previous work has shown that these attributes are related to stronger uptake of reform (Desimone et al., 2005, 2019; Edgerton et al., 2020). Thus, we hypothesize that the stronger the policy features, the more likely it is teachers will alter their instruction in ways supported by the curriculum and PL.

Improving Instruction: Fostering More Ambitious and CR and Less Traditional Instruction

Instructional reform across content areas has called for shifts from a predominant emphasis on traditional approaches (ie, an emphasis on direct instruction, memorization, and procedural tasks) to inclusion of instruction that is both ambitious and CR (International Literacy Association [ILA], 2019; National Council of Teachers of Mathematics [NCTM], 2000). Ambitious practices are those that focus on active learning and conceptual tasks—engaging students in discussion, critical thinking, and real-world applications (Choppin et al., 2020; H. C. Hill et al., 2018; Lampert, 1992; National Mathematics Advisory Panel, 2008). Despite research showing that ambitious teaching is associated with greater student engagement (Turner et al., 2014) and achievement (Le et al., 2009), teachers primarily enact mostly traditional approaches (M. Comstock, A. S. Kenneth et al., 2022; Covay et al., 2015; Desimone & Long, 2010; Desimone et al., 2005). Thus, we need more evidence to guide curriculum and PL interventions to foster more use of ambitious practices.

CR teaching is a paradigm for instruction that uses students' cultural identities as resources for instruction (Delpit, 2006; Gay, 2000, 2010). CR teaching emphasizes teachers' use of pedagogy that builds on students' strengths and supports students to construct knowledge and build social consciousness (Gutstein et al., 1997; Ladson-Billings, 1995, 2009). In a comprehensive review of primarily small-scale and descriptive studies, Aronson and Laughter (2016) found that CR is consistently linked to student outcomes such as increased engagement (Copenhaver, 2001; M. L. Hill, 2009; Nykiel-Herbert, 2010) and improved academic achievement (Gutstein, 2003; Langlie, 2008).

While we know of no randomized, controlled trials of CR instruction, Blazar's (2021) study that randomly assigned White teachers and teachers of color to upper elementary classrooms found substantial and significant effects

Version of Record at: https://doi.org/10.1177/23328584241310429

on student self-efficacy, engagement, attendance, and achievement mediated in part through teachers of color practicing behaviors aligned with CR, such as developing interpersonal relationships with students and individualizing instruction.

The field could benefit from evidence of how to support teachers in using more CR approaches and more empirical evidence from larger samples linking such approaches to student outcomes. Here we focus on student engagement, which has been linked to both achievement and socioemotional outcomes (Markowitz, 2018).

It is from this literature that we derive our research question asking: Does increased use of ambitious and CR instruction relate to greater student engagement, and does traditional instruction relate to less student engagement?

PL's Effects on CR Beliefs and Confidence

How do we elicit changes in instruction that result in more CR and ambitious teaching? Scholars suggest that high-quality PL is a critical lever for changing instructional practice and that it works through increasing teachers' knowledge and shifting their beliefs, which, in turn, leads to changes in their instructional practices (eg, Kennedy, 2016). We ground our hypothesis that PL works through changing confidence and beliefs and that PL provides specific, consistent, authoritative, powerful, stable guidance will help foster teacher beliefs consistent with CR teaching and increase their confidence/self-efficacy in doing so. We do not directly measure knowledge but rely on prior research that links knowledge growth with confidence/self-efficacy (Appleton, 1995; Swackhamer et al., 2009) and changes in teacher practice (Roth et al., 2019).

Previous work suggests that CR-related beliefs, such as holding high expectations for all students, valuing cultural diversity, viewing race as a central component of individual identity, and acknowledging the roles that race and culture play in society, are a precursor to teachers' effective use of these approaches (Gay, 2018; Ladson-Billings, 1995; Seriki & Brown, 2017). Also, teachers' confidence and self-efficacy about engaging in such pedagogy are associated with their self-reported use of CR teaching (Civitillo et al., 2019; Gay, 2010; Zee & Koomen, 2016).

Research suggests that CR-focused PL can influence teacher beliefs and confidence (Fitchett et al., 2012; Parkhouse et al., 2019; Psalti, 2007; Tucker et al., 2005) and teacher practice and student outcomes (Aronson & Laughter, 2016; Parkhouse et al., 2019; Savage et al., 2011). Notably, much of this research is qualitative in nature (Bottiani et al., 2018), which allows us to build on this work by testing these relationships with structural equation modeling and larger samples.

From this literature we derived our research question that asks: Does PL work through bolstering teachers' confidence and beliefs to support more use of ambitious and CR instruction in the classroom?

Integrating these findings across areas, we conduct a comprehensive analysis of how policy attributes of curriculum relate to teachers' CR and ambitious instruction, how PL relates to instruction through teacher beliefs and confidence, and how CR and ambitious instruction relate to student engagement.

Methods

We studied 11 school districts across the country engaged in a PL partnership initiative funded by a prominent education philanthropy organization. The funder required districts to serve 50% or more Black and Latino/a students, English learners, and/or students from low-income families. Districts partnered with external organizations such as PL providers and curriculum developers to enact curriculum-embedded PL in mathematics to promote enhanced student learning, particularly for historically marginalized students. The focus of the curriculum was on ambitious and CR instruction. Thus, this study's findings are derived from a set of equity-oriented racially and linguistically diverse public school districts.

Sample and Data-Collection Procedures

Our sample was comprised of 437 math teachers (many of whom also taught ELA and science) in 153 schools (62% response rate), mostly in grades 6 through 8, in 11 districts. We administered a survey electronically via Qualtrics between April and June of 2021 to teachers in the 11 partnerships. We compensated participating teachers with $60 for completion of the survey, which took approximately 30 minutes. Eighty-seven percent of teachers taught at the middle school level, followed by 9% in high schools and 5% in elementary schools. The sample was racially/ethnically diverse (41% Black or African American, 36% White, 8% Latino/a, 7% Asian, 3% multiracial, $<$1% Indigenous/Native American, and $<$1% Middle Eastern) (see Appendix Table A.1). Teachers were asked to report on their mathematics instruction. Although the findings presented are not generalizable to the broader teacher workforce, the sample is illustrative of public schools serving predominantly marginalized students who are engaging in curriculum adoption and PL around increasing ambitious and CR instruction.

Survey Measures

The validity and reliability of survey items and scales can vary based on the source, development, quality of items, and circumstances of administration (Desimone, 2006; Desimone et al., 2010). We have confidence in our measures for several reasons. We drew on scales from sources that provided reliability and validity data for the scales (eg, Bryk et al., 2010; Cruz et al., 2020; Guyton & Wesche, 2005; Phuntsog, 2001; Siwatu, 2007) and documented concurrent validity through correlations with other key instructional approaches and

Version of Record at: https://doi.org/10.1177/23328584241310429

TABLE 1
Descriptive statistic, teacher survey scales

Factor	N	Mean	SD	Min	Max	Possible range
Scale						
Student engagement	433	2.68	0.54	1.60	4.00	1–4
Traditional instruction	430	3.02	0.56	1.33	4.00	1–4
Ambitious instruction	430	2.56	0.62	1.17	4.00	1–4
CR instruction	427	2.40	0.79	1.00	4.00	1–4
Confidence in CR instruction	432	7.91	1.98	1.75	10.00	0–10
Beliefs about CR instruction	433	4.49	0.95	2.00	6.00	1–6
PL policy attributes						
Specificity	436	4.89	1.09	1.00	6.00	1–6
Consistency	436	4.72	1.04	1.00	6.00	1–6
Authority	435	4.82	1.04	1.00	6.00	1–6
Power	417	3.77	0.85	1.00	6.00	1–6
Curriculum policy attributes						
Specificity	431	4.73	0.93	1.00	6.00	1–6
Consistency	430	4.95	0.92	1.00	6.00	1–6
Authority	434	3.90	0.89	1.00	5.92	1–6
Stability (scale 1 of 2)	417	1.46	0.72	1.00	3.67	1–4
Stability (scale 2 of 2)	433	2.15	0.58	1.00	4.00	1–4

CR, culturally responsive; PL, professional learning

student outcomes (Comstock et al., 2022; Griner & Stewart, 2013; Zee & Koomen, 2016). We engaged in cycles of expert review and cognitive interviews for any items that we adapted or developed to ensure clear understanding of the questions (Desimone & LeFloch, 2004; Fowler, 1995; Rowan et al., 2002), and we conducted a small-scale pilot study. Further, rather than asking teachers to evaluate their own instruction, we used behaviorally oriented questions, which have been shown to have substantial reliability and validity when not asked in an accountability environment, and we uses composites that are better able to distinguish among behaviors (Mayer, 1999).

Further, the notably similar scores across districts (see Appendix Table A.6) and the psychometric properties of our scales (see Appendix Table A.4) support the validity of the data (Burstein, 1995; Stecher & Borko, 2002). At the same time, we acknowledge that instruction, self-efficacy, engagement, and confidence are complex constructs, and we have necessarily narrowed their definition to allow measurement. Thus, interpretation of our findings should be limited to our conception and measurement of these constructs. Next we describe each of the measures used for analysis. Table 1 provides descriptives for the scales, and Appendix Table A.2 lists the items.

PL Policy Attributes Scale. The PL Policy Attributes Scale measures the specificity, consistency, authority, and power of the PL in a teacher's school. Because PL policy interventions were all in their first year of funding and persisted for the duration of the study, we did not measure stability. Each set of items asks teachers to rate their agreement with statements on a six-point scale (completely disagree, mostly disagree, slightly disagree, slightly agree, mostly agree, or completely agree). Items for these scales were based on work from the Center for Standards, Alignment, Instruction and Learning (see, eg, Desimone et al., 2019; Edgerton & Desimone, 2018) and a survey developed and tested in partnership with the School District of Philadelphia (Desimone et al., 2016; K. H. Hill et al., 2023). We created an overall PL policy attributes scale by first creating subscales for each of the attributes by averaging all the items within each subscale. Negative items were reverse coded (for this scale and all others). Then, when adding each subscale, we multiplied by the standardized factor loadings for each subscale (see Appendix Table A.3) to reflect each subscale's contribution to the overarching latent construct PL policy attributes.

Curriculum Policy Attributes Scale. The specificity, consistency, and authority items ask teachers to rate their agreement with statements on a six-point scale (completely disagree, mostly disagree, slightly disagree, slightly agree, mostly agree, or completely agree). Stability items ask teachers to report how long the curriculum has been in their school and district using four timespans (<1 year, about 1–2 years, about 3–5 years, or more than 5 years). Curriculum policy attribute items were based on previously developed and validated surveys (Desimone & Hill, 2017; Desimone et al., 2016; Marsh et al., 2005). We created an

Version of Record at: https://doi.org/10.1177/23328584241310429

overall Curriculum Policy Attributes Scale by averaging items within each subscale and then adding each subscale multiplied by its standardized factor loading to reflect its contribution to the overarching Curriculum Policy Attribute Scale (see Appendix Table A.3).

Beliefs About CR Instruction. The beliefs about CR instruction scale (Cronbach's $\alpha = .715$) measures teachers' beliefs regarding approaches and practices designed to support culturally diverse learners on the same agree/disagree scale as earlier. These items are based on the Culturally Responsive Teaching Outcome Expectancy (CRTOE) Scale and Phuntsog (2001).

Confidence in Culturally Responsive Instruction Scale. The Confidence in Culturally Responsive Instruction Scale (Cronbach's $\alpha = .907$) measures teachers' confidence in adapting instruction and instructional materials to meet the needs of culturally diverse students and to make learning meaningful. Teachers were asked to rate their confidence on statements from a scale from 0 (not confident at all) to 10 (extremely confident). These items are based on the Multicultural Efficacy Scale, the Culturally Responsive Teaching Self-Efficacy (CRTSE) Scale, the CRTOE Scale, the Critical Consciousness Scale, and theory (eg, Freire, 2000; Ladson-Billings, 1995).

Culturally Responsive Instruction Scale. The Culturally Responsive Instruction Scale (Cronbach's $\alpha = .837$) measures the frequency with which teachers report implementing culturally responsive practices in their classroom, such as adapting their instruction to meet the needs of learners from diverse cultural backgrounds or using students' cultural backgrounds to make learning meaningful. Teachers reported how often they engaged in specific activities on a four-point scale (never, a few lessons, about half of all lessons, or most or all lessons). These items are based on the Multicultural Efficacy Scale, the CRTSE Scale, the CRTOE Scale, the Critical Consciousness Scale, and theory (eg, Freire, 2000; Ladson-Billings, 1995).

Ambitious Instruction Scale. The Ambitious Instruction Scale (Cronbach's $\alpha = .828$) measures the frequency with which teachers report incorporating conceptual practices such as small-group work, hands-on activities, and student-led discussions into their instruction. Teachers reported how often they engaged in specific activities on the four-point frequency scale. These items were based on a previously developed and validated survey (Desimone et al., 2016) and RAND Corporation's American Teacher Panel Surveys (Opfer et al., 2021).

Traditional Instruction Scale. The Traditional Instruction Scale (Cronbach's $\alpha = .65$) measures the frequency with which teachers incorporate practices such as direct instruction, call and response, and independent work into their instruction (with the same frequency scale as earlier). These items were based on previously developed and validated surveys (K. H. Hill et al., 2023) and RAND Corporation's American Teacher Panel Surveys (Opfer et al., 2021).

Student Engagement Scale. The Student Engagement Scale (Cronbach's $\alpha = .776$) measures the degree of teacher-reported student engagement during instruction, as evidenced by completing work and asking questions. Teachers were asked to reflect on the behaviors of students in their classes during the previous marking period. Teachers reported the number of students exhibiting specific engagement behaviors on a typical day using a four-point scale (none, a few, about half, or most or all). These items are based on previously developed and validated surveys (Desimone et al., 2016).

Psychometric Properties of Scales

We examined the psychometric properties of each scale using confirmatory factor analysis in a structural equation modeling framework. For each confirmatory factor analysis, we examined the fit statistics (eg, Standardized Root Mean Square Residual [SRMR], Comparative Fit Index [CFI], Tucker-Lewis Index [TLI], and Root Mean Square Error of Approximation [RMSEA]). We checked fit statistics to see if they met the cutoffs reported in the literature while remaining flexible (Hu & Bentler, 1999; Kenny & McCoach, 2003; Lai & Green, 2016; McDonald & Ho, 2002). The models demonstrated acceptable fit when considering all the fit statistics (see Appendix Table A.4) .

To check the reliability of scales, we estimated omega coefficients (ω, which reflect the reliability of a one-factor model—an indication of how well the items work together to measure a single construct) in a factorial framework (see Deng & Chan, 2017). The scales demonstrated acceptable internal consistency, as reported earlier. Then we checked the validity of the scales by examining the standardized factor loadings of items defining each latent variable. The standardized loadings that define a latent variable in the hypothesized model, if appreciable in magnitude, would suggest evidence of good convergent validity (Hair et al., 2021; Kline, 2016; Muthén Bengt, 2002; Raykov, 2011). Most scales had standardized factor loadings of .40 and above, which indicates a meaningful contribution of the item to the unidimensional latent construct (Hair et al., 2021; Kline, 2016; Muthén Bengt, 2002; Raykov, 2011). Appendix Table A.4 provides the psychometric properties of the scales. We also checked for problematic multicollinearity (see Appendix Table A.5 for a correlation analysis of the latent variables) and found no correlations higher than .6. While there are debates about the appropriate cutoff for multicollinearity, .7 to .9 and higher are

Version of Record at: https://doi.org/10.1177/23328584241310429

generally thought to be in the problematic range (Yoo et al., 2014).

Data Analytic Approach

We aggregated teachers across districts, given the nature of our sample. As shown in Appendix Table A.6, the number of teachers within each district does not allow for nesting (Hox, 2010). Further, the means for each of our key latent variables are notably similar across districts, suggesting that no one district's responses are driving the results (although we do have considerable variability in numbers of respondents across districts).

We used structural equation modeling (SEM) to model the relationships we are testing. In SEM, researchers apply theory to hypothesize relationships between latent variables (measures of a specific construct) by designing and specifying a priori structural equation models (Bollen, 2014; Bollen & Pearl, 2013; Kline, 2016, 2023; Morgan, 2013). Researchers justify the structural paths leading from one latent variable to another with theory, which represents the study's hypotheses. SEM allows researchers to model simultaneously the direct and moderated paths of multiple relationships, as posited by theoretical notions (Gerbing & Anderson, 1988).

For each of our structural equation models, we examined model fit to evaluate the overall structure using acceptable thresholds reported in the literature: RMSEA $<$ .06, CFI $>$ .90, TLI $>$ .90, and SRMR $<$ .06 (Hu & Bentler, 1999; Kenny & McCoach, 2003; Lai & Green, 2016; McDonald & Ho, 2002). Then we reviewed standardized parameter coefficients and amount of variance explained (eg, R^2 statistics) for each model, only in cases when model fit was deemed acceptable. R^2 indicates the percent of variability in the variable that the model explains, and its associated P values indicate whether the percent of variation explained is significant. Small, medium, and large effects follow guidelines for SEM and Cohen's d, where .2 is small, .5 is medium, and .8 is large (Gomer et al., 2018).

Structural Equation Models. We specified two structural equation models. Our first model tests our hypotheses that (1) teachers who experience curriculum and PL with desirable policy features (e.g., specificity, consistency, authority, power, and stability) use more ambitious and CR instruction and less traditional instruction, (2) curriculum and PL have separate and interactive relationships with instruction, and (3) ambitious and CR instruction have a positive relationship with student engagement, whereas traditional instruction has a negative one (see Figure 1).

In our second structural equation model, we test the hypothesis that the policy attributes of PL indirectly relate to instruction by working through teachers' beliefs and confidence (see Figure 2).

Statistical Software and Missing Data. We used Mplus, version 8.3, to compute statistical analyses. For analyses, we estimated parameters of structural equation models by using full-information maximum likelihood. Although missing data was not a concern, in that we had $<$5%, full-information maximum likelihood uses all the data for estimation, producing unbiased estimates (Asparouhov & Muthén, 2019; Enders, 2003; Enders & Bandalos, 2001).

Results

Prior to interpreting standardized coefficients, we checked for acceptable model fit. We were unable to assess the fit of the first model because statistics (e.g., SRMR, CFI/TLI, and RMSEA) have not been developed for latent variable interaction models. The second model had an SRMR = .07, CFI/TLI = .90/.89, and RMSEA = .045 (range, .042–.049), which demonstrated acceptable fit based on our a priori thresholds, suggesting that the relationships set forth in our hypothesized model predicted the associations among variables.

Structural Equation Model 1: How the Policy Attributes of PL and Curriculum Relate Separately and Interactively to Ambitious, CR, and Traditional Instruction and, In Turn, Student Engagement

The results, shown in Figure 3, support our first hypothesis that stronger policy features of curriculum and PL are related to teachers' use of ambitious (β=.104, P=.014; β=.161, P=.003) and CR instruction (β=.158, P=.019; β=.205, P=.013) but not traditional instruction (β=.076, P=.113; β=0, p=.710). Here the beta coefficient indicates how many standard deviation (SD) units are associated with a 1-unit increase in the independent variable. For example, a 1 SD increase in policy features of PL is related to .161 SD unit increase (P=.003) in ambitious instruction, after controlling for CR and traditional instruction.

Findings also support our hypothesis that PL and curriculum work together to bolster ambitious (β=.085, P=.001) and CR instruction (β=.099, P=.014). The coefficient for traditional instruction (β=.059, P=.049) was marginally statistically significant. The effect sizes for CR (r^2=.101, P=.001) and ambitious instruction (r^2=.165, P=.001) indicate that curriculum, PL, and their interaction described 10.1% of the variability in CR instruction and 16.5% of the variability in ambitious instruction, a medium and large effect size, respectively (see Table 2).

We hypothesized that CR instruction and ambitious instruction were related to increased student engagement and that traditional instruction was related to decreased engagement, but in our models, only the use of ambitious instruction was associated with more student engagement (β=.488, $P<$.000), explaining 14.5% of the variability (r^2=.145, P=.006). CR instruction was significantly

Version of Record at: https://doi.org/10.1177/23328584241310429

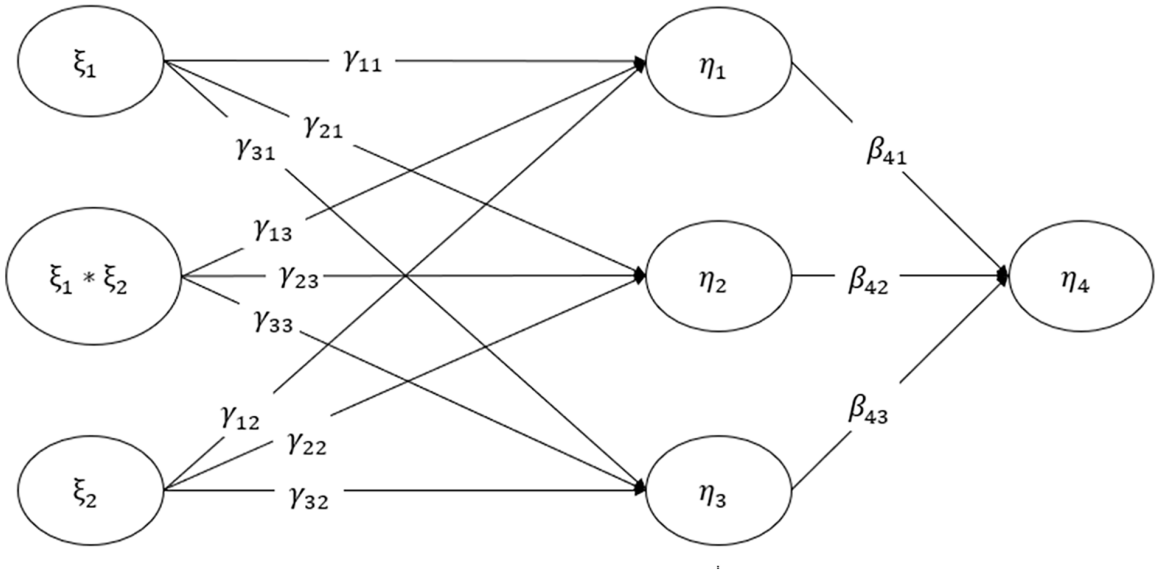

FIGURE 1. *Structural Equation Model 1.*

Note. Structural Equation Model 1 is represented by the following series of equations:

$$\eta_1 = \alpha_1 + \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \gamma_{13}\xi_1\xi_2 + \zeta_1$$

$$\eta_2 = \alpha_2 + \gamma_{21}\xi_1 + \gamma_{22}\xi_2 + \gamma_{23}\xi_1\xi_2 + \zeta_2$$

$$\eta_3 = \alpha_3 + \gamma_{31}\xi_1 + \gamma_{32}\xi_2 + \gamma_{33}\xi_1\xi_2 + \zeta_3$$

$$\eta_4 = \alpha_4 + \beta_{41}(\eta_1) + \beta_{42}(\eta_2) + \beta_{43}(\eta_3) + \zeta_4$$

where η = latent endogenous variable, α = intercept, γ = regression coefficient of exogenous-to-endogenous variables, ξ = latent exogenous variable, η_{x1} = latent variable of CR instruction, β = regression coefficient of endogenous-to-endogenous variables, and ζ = residual model error.

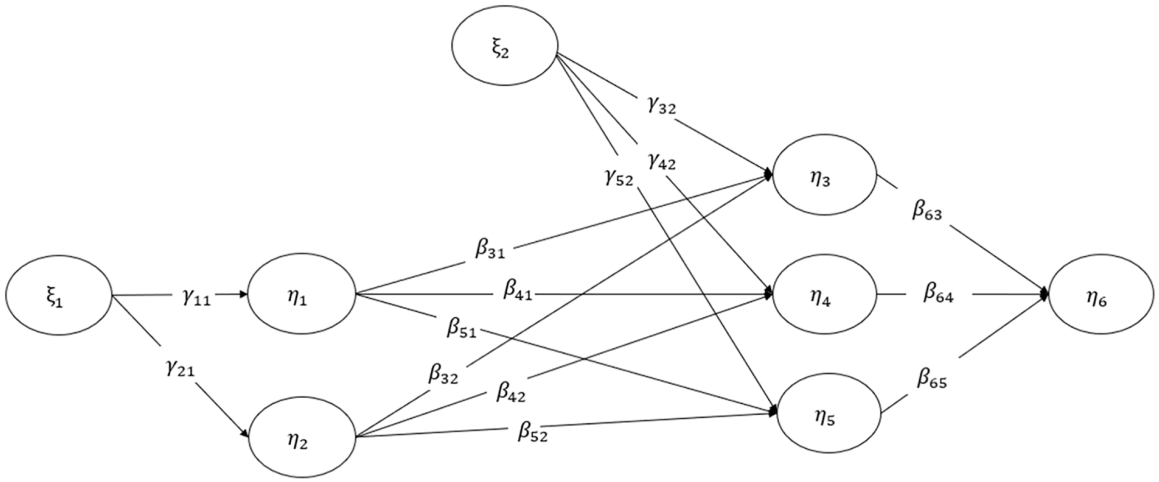

FIGURE 2. *Structural Equation Model 2.*

Note. Structural Equation Model 2 is represented by the following equations:

$$\eta_1 = \alpha_1 + \gamma_{11}(\xi_1) + \zeta_1$$

$$\eta_2 = \alpha_2 + \gamma_{21}(\xi_1) + \zeta_2$$

$$\eta_3 = \alpha_3 + \gamma_{32}(\xi_2) + \beta_{31}(\eta_1) + \beta_{32}(\eta_2) + \zeta_3$$

$$\eta_4 = \alpha_4 + \gamma_{42}(\xi_2) + \beta_{41}(\eta_1) + \beta_{42}(\eta_2) + \zeta_4$$

$$\eta_5 = \alpha_5 + \gamma_{52}(\xi_2) + \beta_{51}(\eta_1) + \beta_{52}(\eta_2) + \zeta_5$$

$$\eta_6 = \alpha_6 + \beta_{63}(\eta_3) + \beta_{64}(\eta_4) + \beta_{65}(\eta_4) + \zeta_6$$

where η = latent endogenous variable, α = intercept, γ = regression coefficient of exogenous-to-endogenous variables, ξ = latent exogenous variable, η_{x1} = latent variable of CR instruction, β = regression coefficient of endogenous-to-endogenous variables, and ζ = residual model error.

Version of Record at: https://doi.org/10.1177/23328584241310429

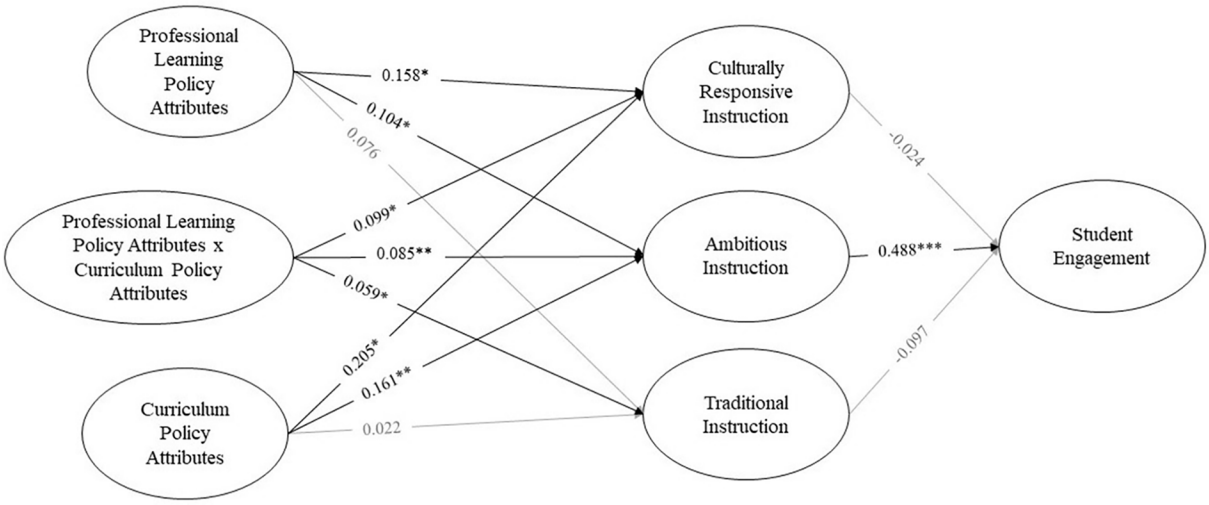

FIGURE 3. *Structural Equation Model 1.*
*$P < .050$; **$P < .010$; ***$P < 0.001$.

TABLE 2
Effect sizes.

	Model 1		Model 2	
Factor	r^2	*P* value	r^2	*P* value
Belief in CRP	—	—	.052	.045
Confidence in CRP	—	—	.138	.000
CR instruction	.101	.001	.228	.000
Ambitious instruction	.165	.001	.157	.000
Traditional instruction	.038	.249	.034	.143
Student engagement	.145	.006	.136	.002

CRP, culturally responsive pedagogy; CR, culturally responsive

associated with an increase in student engagement when ambitious instruction was not in the model (β = .22, $P < .009$). We interpret this to suggest that much CR instruction is done in the context of ambitious teaching, and when it is not, it may not engage students to its fullest potential.

Structural Equation Model 2: Extent to Which PL Works Through Changing Teachers' Confidence and Beliefs

Next, we wanted to better understand how PL works—specifically, we tested our hypothesis that the attributes of PL work through bolstering teachers' beliefs and confidence in CR instruction. As with Model 1, we also examined the direct paths from curriculum to instruction and from instruction to student engagement.

Similar to Model 1, curriculum policy attributes were associated with an increase in CR instruction (β = .135, $P = .010$) and ambitious instruction (β = .293, $P < .000$) but not traditional instruction (β = .056, $P = .397$). We also found that the stronger the policy features of the PL, the more likely teachers were to believe in and have confidence in CR approaches. Specifically, the results in Figure 4 show that PL was associated with increased teachers' beliefs (β = .228, $P < .000$) and confidence in CR instruction (β = .371, $P < .000$), that PL explained 13.8% of the variability in teachers' confidence in CR instruction ($r^2 = .138$, $P < .000$), a medium effect size, and that PL explained 5.2% of the variability in teachers' belief in CR instruction ($r^2 = .052$, $P = .045$), a small effect size.

This model also showed that CR beliefs were not related to using more ambitious (β = −.073, $P = .267$) or traditional instruction (β = −.149, $P = 0.051$) but were related to more frequent CR instruction (β = .130, $P = .041$). Confidence in CR instruction was associated with an increase in ambitious (β = .215, $P < .000$) and CR instruction (β = .342, $P < .000$) but not traditional instruction (β = .158, $P = .021$, but the effect size (r^2) was not significant, indicating that none of the variability in traditional instruction was explained by our model.

Model 2 explained variability in ambitious and CR instruction but was only marginally significant in explaining traditional instruction; this may reflect how entrenched traditional instruction is. Model 2 explained 15.7% of the variance in ambitious instruction ($r^2 = .157$, $P < .00$) and 22.8% of the variance in CR instruction ($r^2 = .228$, $P < .000$), both considered relatively large effect sizes.

Consistent with findings from Model 1, ambitious instruction was associated with *increased* student engagement (β = .514, $P < .000$), whereas neither of the relationships of CR instruction (β = −.116, $P = .132$) and traditional instruction (β = −.190, $P = .050$) with engagement were statistically significant. Ambitious instruction explained 13.6% of the variance in student engagement ($r^2 = .136$, $P = .002$), a medium effect size.

Version of Record at: https://doi.org/10.1177/23328584241310429

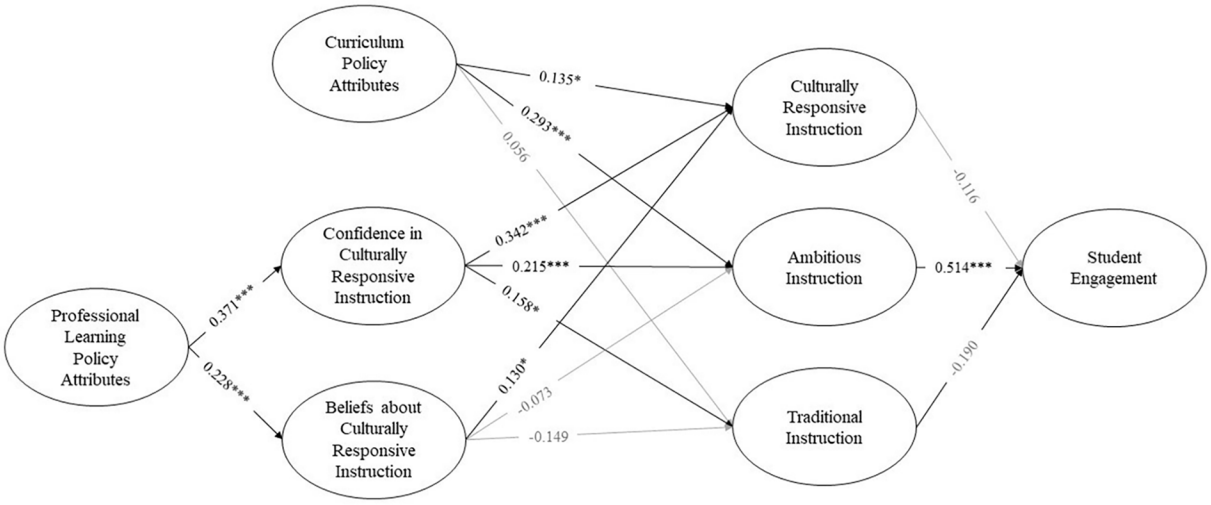

FIGURE 4. *Structural Equation Model 2.*
*$P < .050$; **$P < .010$; ***$P < 0.001$.

We also tested a model (not shown) that added to Model 2 the direct relationships of PL with instruction, but those paths were not statistically significant, providing more evidence that PL works through teachers' confidence and beliefs.

Limitations

Our findings should be interpreted considering both their strengths and limitations. Our sample is comprised primarily of middle school mathematics teachers. Future research might examine whether our results are similar for other grades and subject areas. Also, we relied on survey data, which can be subject to respondent bias and social desirability, especially surveys about cultural competence (Larson & Bradshaw, 2017). To alleviate some of this bias, we used behavior-based measures of teacher instruction, which prior work has shown to be a valid and reliable way of measuring teacher practice (Mayer, 1999). Although we partly addressed this bias by conducting reliability and validity psychometric testing of the scales, we acknowledge that teachers' self-reports may not fully or accurately reflect their classroom practice (see, eg, Parkhouse et al. [2019] regarding CR teaching specifically). Future studies might build on ours by using complementary measures such as classroom observations. Similarly, our outcomes are instruction and student engagement, not student achievement. We rely on the literature that links use of ambitious and CR instruction and student engagement (eg, Cooper, 2014; Goldin et al., 2019; Marks, 2000) to increased learning (eg, Cherfas et al., 2021; Portes et al., 2018) and student engagement with student learning (Markowitz, 2018; Wang & Holcombe, 2010). Further, while a strength of the study is our measuring a set of key policy features of the curriculum and PL, we are not able to identify the independent contribution of each of the features; designing studies that isolate features would be a future next step that would be helpful in understanding what may be driving effects.

We do not use an experimental design, which would allow causal claims. While SEM is considered by many scholars to be an appropriate method for investigating causal relationships using nonexperimental data (Bollen, 2014; Heckman, 2005; McDonald, 2004; Meehl & Waller, 2002; Morgan, 2013), we adopt a conservative approach with our cross-sectional data and interpret our findings as relational. SEM is especially well suited to model complex interrelated relationships (Bollen & Pearl, 2013; Gerbing & Anderson, 1988; Tarka, 2018) such as the ones we examined in this analysis. Despite the caveats that alternative models also may fit the data (Cliff, 1983), there may be omitted variable bias (Bollen, 1989), and our results rely on the accurateness of a priori theoretical models grounded in previous research (Bollen & Long, 1993). This approach has the potential to provide valuable insights into making sense of complexity (Macallum, 2003). As such, our work offers important insights into understanding how curriculum, PL, beliefs, and confidence operate as a system to influence instruction and student engagement; prior research has rarely modeled these complex relationships simultaneously.

Discussion

How Does Curriculum and PL with Specific High-Quality Features Relate to Teachers' Use of Ambitious, CR, and Traditional Instruction? Are Efforts More Successful When a New Curriculum Is Coupled with Aligned PL?

Curriculum and PL are arguably the two most common leverage points for school and instructional improvement (eg, Penuel et al., 2007). We explore them in the context of

Version of Record at: https://doi.org/10.1177/23328584241310429

ambitious and CR initiatives and examine whether they work better together and the extent to which a set of strong policy features is associated with instructional change.

Our findings that stronger curriculum and aligned PL attributes (ie, specificity, consistency, authority [buy-in], and power [rewards and sanctions]) were related to teachers' increased use of ambitious and CR instruction are consistent with previous findings of the importance of specificity (Stornaiuolo et al., 2023), authority (Pak et al., 2021), and power (Nichols et al., 2021) in the uptake of reform. Our finding that these attributes did not relate to teachers' use of traditional (procedural and transactional) approaches is consistent with the enduring nature of teacher-centered instructional approaches (eg, Lefstein, 2008; Litke, 2020) and research that shows that teachers' responses to improvement initiatives are typically to add on to their instructional routines without omitting previous practices (eg, Cohen, 1990). Also, as teachers learn new ideas and approaches in PL, we might expect them to increase their use of multiple types of practices including traditional as they build their expertise in using different approaches in diverse ways. This complicates the interpretation of instructional change, when teachers layer on different approaches rather than trade off one for another. Still, *it is encouraging that carefully selected high-quality curricula and purposefully crafted PL focused on CR and ambitious instruction are associated with increased use of these approaches in the classroom.*

Our finding that combining aligned curriculum and PL bolsters the use of ambitious and CR instruction even more than either alone supports the idea that aligning PL with curricula plays a key role in motivating instructional improvement (Polikoff, 2021) and that found curricular-aligned PL is more effective than either alone (Lynch et al., 2019; Pellegrini et al., 2021). Our results from multiple districts and schools across the country build on earlier work that studied a particular intervention and found that PL was more effective when it is explicitly linked to curricula (Penuel et al., 2007, 2011; Taylor et al., 2015) and a review of randomized, controlled trials in PL that found that PL was more effective when closely linked to teachers' lessons and curricula (Desimone & Garet, 2015). Our contribution to this literature on single interventions or syntheses of single interventions is that we are able to explicitly measure in the same analysis features of PL and curriculum interventions in multiple districts across the country to model both direct and interactive relationships with instruction.

Our results provide substantial evidence that when PL and curriculum are targeted toward ambitious and CR instruction, provide clear, specific direction to teachers, are aligned with each other, have the support and buy-in of teachers, and operate in a policy system that has incentives for enacting the curriculum and the approaches taught in the PL, teachers are more likely to use ambitious and CR instruction in the classroom. This set of policy features can serve as a roadmap to districts to guide their design of PL and associated policy systems to support instructional improvement.

Does Increased Use of Ambitious and CR Instruction Relate to Greater Student Engagement, and Does Traditional Instruction Decrease Engagement?

Here we build on previous qualitative and case study work (eg, Duke et al., 2021; Irvine, 2020; Turner et al., 2014) to respond to the call for complementary larger-scale evidence on the relationship between ambitious and CR instruction and student outcomes (Sleeter, 2012).

We expected more frequent enactment of ambitious and CR instruction to result in higher levels of reported student engagement (eg, Feger, 2006; Gay, 2002; Ginsberg, 2015; Harbour et al., 2015), but in our analysis, only ambitious instruction increased student engagement. Our findings are similar to those of Byrd's (2016) national survey of 315 middle and high school students that showed that ambitious instruction (labeled *constructivist*) was related to student interest and belonging but that practices related to cultural and critical consciousness were not. A careful look at our CR measures might help explain this. Our CR measured included items that ask teachers about examining materials for stereotypes and adapting materials to address bias. These types of practices fall within CR instruction, but they do not imply a particular pedagogy. They may result in students feeling valued and respected (Gay, 2021; Ladson-Billings, 1995), bolstering their perception of themselves as capable students (Souryasack & Lee, 2007) and increasing their confidence in school (Hubert, 2014), but this may take a while to translate into engagement. Furthermore, such changes may not be captured by our measure of engagement, which included items such as completing homework, persevering through challenges, asking questions, and working independently.

Additionally, our results here are consistent with an idea that so far has not been assessed empirically, that CR instruction is most effective when conducted in the context of ambitious instruction. In our original conception, we drew on considerable literature that distinguishes ambitious approaches that focus on cognitive demand, challenging content, and active learning (eg, Kraft & Hill, 2020; Smith et al., 2005, 2007) from CR instruction that individualizes instruction, draws on students' funds of knowledge, and includes an emphasis on social and political critique (eg, Abdulrahim & Orosco, 2020; Bottiani et al., 2018; Howard & Rodriguez-Minkoff, 2017).

In separate models, *each had a positive and significant relationship with student engagement, but when both were in the same model, only ambitious instruction was significantly related to student engagement. We suspect this is due in part to the interactive relationship between CRT and ambitious instruction*. Many teachers use ambitious and CRT strategies in concert; our descriptive statistics bear this out. Ambitious

Version of Record at: https://doi.org/10.1177/23328584241310429

and CR instruction were statistically significantly correlated at .56; specifically, 56% of teachers who reported very frequent use of ambitious instruction also reported very frequent use of CR instruction. Taken together, these findings suggest the potential for districts to integrate ambitious and CR strategies for maximum benefit to students. Further research might explore if and how ambitious and CR approaches can and should be combined for maximum student engagement and learning. Understanding how to leverage and interact CR and ambitious strategies is a challenge for the field.

We expected that an emphasis on traditional instruction, which uses teacher-led activities, memorization, and procedural tasks, would reduce student engagement (see Markowitz, 2018), but in fact, in our models there was no statistically significant relationship between traditional instruction and student engagement. One interpretation of these results is that traditional instruction is so entrenched in US public schools that student engagement levels are impervious to it. We are encouraged by the findings that ambitious approaches are related to increased engagement, however.

Does PL Work Through Bolstering Teachers' Beliefs and Confidence to Support More Use of Ambitious and CR Instruction in the Classroom?

Building on previous work demonstrating a link between PL and more ambitious and CR teaching (Horn & Garner, 2022), we were interested in whether the relationship of PL with the use of ambitious and CR teaching was mediated by teachers' confidence and beliefs, and we found that it was. PL did not have a direct relationship with instruction. Although our first model shows that teachers who participate in PL that is specific, consistent, authoritative, powerful, and stable use more ambitious and CR instruction, our second model shows that these relationships are explained in part by teacher beliefs and confidence.

Thus, we provide empirical evidence that *the relationship between PL and CR instruction is mediated by teachers' beliefs and confidence related to CR instruction*. PL was associated with stronger teachers' beliefs in the value of CR practice, that it is essential for an inclusive classroom, that incorporating social or political issues and encouraging students to be politically active are appropriate activities, and that different students may learn from various teaching methods. Also, PL was associated with higher teacher confidence in using examples from students' everyday lives, in integrating social or political issues into class discussions, in supporting students in social or political action, and in uncovering their own bias in their teaching, and this was associated with more CR instruction. Many frameworks conceptualizing how PL works posit that it works through bolstering confidence and beliefs (Borko, 2004; Desimone, 2009; Kennedy, 2016). We add to the literature by formally modeling these mediating relationships between PL and instruction. We also found that CR confidence is related to more traditional teaching, although much less so than for ambitious and CR instruction. This could reflect the phenomenon we discussed earlier, that developing expertise in a specific area enables a teacher to use multiple approaches to good effect.

Another key finding from our study is that *confidence matters more than beliefs for fostering change in instruction*. Confidence in CR instruction was associated with more use of both ambitious and CR instruction. Belief in CR instruction was associated with more use of CR instruction (β = .130), but by less than half the relationships with confidence (β = .342). This is in line with previous work that while beliefs can be critical precursors to changes in practice (Spillane et al., 2018), beliefs do not always translate directly into practice (Guerra & Wubbena, 2017) and that teachers need to feel self-efficacious to put their beliefs into practice. Our analysis demonstrates that confidence in using strategies consistent with CR beliefs is a precursor to more use of CR strategies in the classroom. An implication of this finding is that districts should attend explicitly to supporting teachers in developing their confidence and beliefs in CR instruction.

Conclusion

Our study advances the field by explicitly measuring features of PL and curriculum in multiple districts across the country to model both direct and interactive relationships with instruction. Our results provide evidence, from a sample of 437 teachers in 153 schools across 11 districts serving predominantly Black and Brown children, that when PL and curriculum are targeted toward ambitious and CR instruction, provide clear, specific direction to teachers, are aligned with each other, have the support and buy-in of teachers, and operate in a policy system that has incentives for enacting the curriculum and the approaches taught in the PL, teachers are significantly more likely to use ambitious and CR instruction in the classroom. Further, we found that these relationships are both direct and interactive. That is, PL and curriculum each have independent direct relationships with more use of ambitious and CR instruction in the classroom, and when provided together, they relate even more strongly to teachers' use of ambitious and CR instruction.

We showed that the PL worked through changing teacher beliefs and confidence; we provided evidence that specific, authoritative, consistent, and powerful PL can help shift teacher beliefs to consider CR instruction as appropriate and useful and develop teacher confidence in employing CR strategies in the classroom. These changes in beliefs and confidence, in turn, were associated with more use of CR and ambitious instruction in the classroom. Finally, we found that ambitious instruction—where students work on extended activities; apply knowledge to new situations; critique, synthesize, and evaluate; lead discussions and presentations; and self-reflect—significantly and

Version of Record at: https://doi.org/10.1177/23328584241310429

substantially increased student engagement. Our findings apply to our middle school mathematics sample but may hold promise for other grades and subjects.

We are encouraged by these findings. In our quest to provide engaging, challenging, an CR instruction to all students across the country, our study offers evidence across schools and districts that this can be achieved. We hope that this study contributes to other work that is helping to identify ways that schools and districts can shape their learning and curriculum to better serve all our nation's students.

Appendix

TABLE A.1
Characteristics of survey respondents

Teacher characteristic	No. of respondents	Percent of sample
Grade level taught by teacher		
Elementary school and middle school	21	4.8%
Middle school only	381	87.2%
High school and middle school	35	8.7%
Subject taught by teacher[a]		
ELA	14	13.0%
Math	386	85.6%
Not indicated	37	12.1%
Race/ethnicity of teacher[b]		
Asian American/AAPI	29	6.6%
Black/African American	178	40.7%
Indigenous/Native American	6	1.4%
Latinx	36	8.2%
Middle Eastern	5	1.1%
Multiracial	12	2.7%
White	159	36.4%
Prefer not to say	35	8.0%

ELA, English language arts; AAPI, Asian American Pacific Islander
[a]Teachers who reported teaching ELA likely taught math in the prior year; teachers who did not indicate that they taught math answered the math-related questions, so we included them in the sample rather than omitting their responses because of the missing data.
[b]Teachers selected "All that apply" when selecting race/ethnicity, and we used the total sample as the denominator (437), so the percentages total to >100. In comparison with our sample, the 2017–2018 US teacher labor force was 79% White, 7% Black, 9% Hispanic, 2% Asian, 1% American Indian/Alaska Native, and 2% multiracial (National Center for Education Statistics, 2019).

TABLE A.2
Survey items used for analysis by scale

Survey question	Response options
Beliefs about culturally responsive instruction (standardized loadings in parentheses)	
Please mark the extent to which you agree or disagree with each of the following statements:	Completely disagree
1. Culturally responsive practice undermines classroom unity by emphasizing cultural differences. (.30)[a]	Mostly disagree
2. Culturally responsive practice is essential for creating an inclusive classroom environment. (.44)	Slightly disagree
3. Regardless of cultural differences, all children learn from the same teaching method. (.49)[a]	Slightly agree
4. Teachers should not incorporate social or political issues into their instruction. (.54)[a]	Mostly agree
5. A color-blind approach to teaching is effective for ensuring respect for all culturally diverse students. (.53)[a]	Completely agree
6. Encouraging students to be active in social or political causes is an important role for a teacher. (.72)	
Confidence in culturally responsive instruction (standardized loadings in parentheses)	
Please rate your confidence in the following areas on a scale from 0 (not confident at all) to 10 (extremely confident):	0–10
1. Explaining new concepts using examples that are taken from my students' everyday lives. (.42)	
2. Integrating social or political issues into class discussions or assignments. (.62)	
3. Uncovering my own implicit biases in my teaching practice. (.92)	
4. Supporting my students to be active in social or political causes (.93)	

(continued)

Version of Record at: https://doi.org/10.1177/23328584241310429

TABLE A.2 (CONTINUED)

Survey question	Response options
Culturally responsive instruction	
In the previous marking period for [math or English language arts (ELA)], about how often did you engage in each of the following activities?	Never A few lessons About half of all lessons Most or all lessons
1. Adapt instructional methods to meet the needs of learners from diverse cultural backgrounds. (.59)	
2. Analyze instructional materials for potential stereotypical and/or prejudicial content. (.62)	
3. Develop activities designed to increase the self-confidence of students from different cultural backgrounds. (.62)	
4. Use my students' cultural background to help make learning meaningful. (.73)	
5. Identify cultural biases in textbooks or other instructional materials. (.79)	
6. Use the interests of my students to make learning meaningful for them. (.80)	
7. Revise instructional materials to include a better representation of cultural groups. (.81)	
8. Design a lesson that shows how different cultural groups use the material. (.82)	
Ambitious instruction	
In the previous marking period for [math or ELA], about how often did you have students engage in the following activities?	Never, A few lessons, About half of all lessons, Most or all lessons
1. Work on extended learning activities (eg, projects or portfolios). (.54)	
2. Apply their knowledge to new situations, concepts, or problems. (.55)	
3. Critique, evaluate, or synthesize. (.60)	
4. Participate in student-led discussions. (.68)	
5. Present work to the class. (.70)	
6. Participate in self-reflection. (.75)	
Traditional instruction	
In the previous marking period for [math or ELA], about how often did you have students engage in the following activities?	Never A few lessons About half of all lessons Most or all lessons
1. Learn or practice basic facts, concepts, and procedures related to a topic. (.49)	
2. Work independently. (.49)	
3. Answer multiple-choice, fill-in-the-blank, or true/false questions (eg, worksheets, quizzes, tests, or warm-ups). (.53)	
4. Receive direct instruction. (.53)	
5. Engage in call and response. (.55)	
6. Take notes from lectures or the textbook. (.64)	
Student engagement	
In the previous marking period, on a typical day, about how many of the students in your [math or ELA] class(es) do each of the following?	None, A few About half Most or all
1. Complete their assigned work. (.45)	
2. Give up if they find that their schoolwork in your class(es) challenging. (.51)	
3. Ask you questions when they need help. (.62)	
4. Are engaged in their learning. (.75)	
5. Work independently to complete tasks without your help. (.83)	
Professional learning policy attribute—specificity	
Please mark the extent to which you agree or disagree with each of the following statements: My professional learning related to [math or ELA] . . .	Completely disagree Mostly disagree Slightly disagree Slightly agree Mostly agree Completely agree
1. has clear objectives and goals for my learning.	
2. presents information that is clearly and explicitly linked to the curriculum.	
3. presents information that is clearly and explicitly linked to my state's content standards.	
4. provides explicit guidance on how to integrate the curriculum into my instruction.	
Professional learning policy attribute—consistency	
Please mark the extent to which you agree or disagree with each of the following statements: The professional learning activities for [math or ELA] are . . .	Completely disagree Mostly disagree Slightly disagree Slightly agree Mostly agree Completely agree
1. integrated into/linked with my daily lessons/curricula.	
2. aligned with my school's mission and goals.	
3. consistent with district policies (such as state standardized testing and content standards).	
4. consistent with feedback from observations of my teaching.	
Professional learning policy attribute—authority	
Please mark the extent to which you agree or disagree with each of the following statements about professional learning for [math or ELA]:	Completely disagree Mostly disagree Slightly disagree Slightly agree Mostly agree Completely agree
1. I believe that the strategies I learned will improve my [math or ELA] instruction.	
2. The strategies taught are consistent with my values and beliefs about instruction.	
3. I feel supported by my school leadership in implementing what I learned into my classroom practice.	
4. I feel supported by my district in implementing what I learned into my classroom practice.	
5. My participation in professional learning on [math or ELA] can lead to better achievement outcomes for students.	
Professional learning policy attribute—power	
Please mark the extent to which you agree or disagree with each of the following statements:	Completely disagree Mostly disagree Slightly disagree Slightly agree Mostly agree Completely agree
1. There are incentives to encourage teachers to participate in professional learning related to [math or ELA].	
2. There are negative repercussions for teachers who do not attend professional learning for [math or ELA].[a]	
3. Teachers who implement [math or ELA] poorly will receive a lower evaluation rating.[a]	
4. Teachers are recognized for exemplary [math or ELA] teaching.	
5. I feel pressure to teach to my state's content standards for [math or ELA].[a]	
6. There are negative repercussions for teachers whose students perform poorly on my state's standardized achievement test.[a]	

(continued)

Version of Record at: https://doi.org/10.1177/23328584241310429

TABLE A.2 (CONTINUED)

Survey question	Response options
Curriculum policy attributes—specificity	
Please mark the extent to which you agree or disagree with each of the following statements: [Math or ELA] . . .	Completely disagree
1. clearly indicates the [math or ELA] content I should teach.	Mostly disagree
2. provides detailed guidance on [math or ELA] objectives.	Slightly disagree
3. provides the sequence in which topics are covered in [math or ELA].	Slightly agree
4. sets the pace for covering topics in [math or ELA].	Mostly agree
5. dictates the teaching methods and strategies I should use with my students in my [math or ELA] class(es).	Completely agree
6. provides criteria for grading students in [math or ELA].	
7. recommends evaluation and assessment activities in [math or ELA].	
8. recommends books and other materials to support my [math or ELA] instruction.	
Curriculum policy attributes—consistency	
Please mark the extent to which you agree or disagree with each of the following statements: [Math or ELA] . . .	Completely disagree
1. promotes consistency of instruction among [math or ELA] classes at the same grade level.	Mostly disagree
2. promotes continuity of [math or ELA] instruction between grades.	Slightly disagree
3. is aligned with my state's [math or ELA] content standards.	Slightly agree
4. is aligned with my state's [math or ELA] assessments.	Mostly agree
5. is aligned with my district or school's [math or ELA] formative and/or summative assessments.	Completely agree
Curriculum policy attributes—authority	
Please mark the extent to which you agree or disagree with each of the following statements:	Completely disagree
1. I have the training I need to implement [math or ELA].	Mostly disagree
2. I have the resources I need to implement [math or ELA].	Slightly disagree
3. I have adequate time to implement [math or ELA].	Slightly agree
4. I need to supplement [math or ELA] to meet the needs of my students.[a]	Mostly agree
5. I believe that [math or ELA] can lead to better achievement outcomes for students.	Completely agree
6. [Math and ELA] are too inflexible for me to effectively teach my students.	
7. [Math and ELA] include more content than can be covered adequately in the school year.[a]	
8. [Math and ELA] are too rigorous for most of the students I teach.[a]	
9. [Math and ELA] help me prepare my students for state standardized tests.[a]	
10. [Math and ELA] exclude important content that students should learn.[a]	
11. [Math and ELA] appropriately address the needs of students who are designated as English learners.	
12. [Math and ELA] appropriately address the needs of students who have individualized education programs or the equivalent.	
Curriculum policy attributes—stability	
Approximately how many years have you personally used [math or ELA] in your classroom?	Less than a year
Approximately how many years has [math or ELA] been required by your school and/or district?	About 1-2 years
	About 3–5 years
How much longer do you think [math or ELA] will be required by your school and/or district?	More than 5 years
To what extent do you consider each of the following factors a challenge to implementing [math or ELA] in your classroom?	Not a challenge,
1. The way instruction was organized this year (e.g., virtual, hybrid, etc.).	A slight challenge,
2. Frequent changes in district initiatives.	A moderate challenge,
3. Turnover in district leadership (e.g., the superintendent).	A great challenge
4. Frequent changes in school priorities.	
5. Turnover in school leadership.	
6. Teacher turnover.	
7. Teacher absenteeism.	
8. Students transferring into or out of my class.	
9. Student absenteeism.	
10. Student tardiness.	

Notes. Standardized loadings are in parentheses. See Table A.3 for subscale loadings for PL and curriculum.
[a]These items were reverse coded for analysis.

TABLE A.3

Standardized factor loadings for professional learning policy attributes and curriculum policy attributes

Policy features subscales	Standardized loadings:Professional learning policy attributes	Standardized loadings: Curriculum policy attributes
Specificity	.33	.22
Consistency	.45	.25
Authority	.87	.68
Power	.89	—
Stability	—	.80
Curriculum implementation context	—	.80

Version of Record at: https://doi.org/10.1177/23328584241310429

TABLE A.4
Psychometric properties of scales

Latent variable	Omega	Standardized loading	SRMR	CFI/TLI	RMSEA
Professional learning policy attributes	.75*	.33–.90*	.03	.99/.99	.049 (.001–.092)
Curriculum policy attributes	.81*	.22–.80*	.03	.98/.96	.069 (.039–.101)
Beliefs about culturally responsive instruction	.71*	.301–.72*	.05	.98/.96	.053 (.011–.092)
Confidence in culturally responsive instruction	.88*	.42–.93*	.03	.99/.95	.126 (.073–.187)
Culturally responsive instruction	.92*	.59–.82*	.03	.97/.96	.068 (.053–.084)
Ambitious instruction	.80*	.54–.75*	.04	.95/.90	.106 (.077–.136)
Traditional instruction	.68*	.49–.64*	.02	1.00/1.00	.001 (.001–.040)
Student engagement	.77*	.45–.83*	.03	.98/.94	.086 (.046–.130)

SRMR, Standardized Root Mean Square Residual; CFI/TLI, Comparative Fit Index/Tucker-Lewis Index; RMSEA, Root Mean Square Error of Approximation
Note. For the RMSEA, the first number is the estimate, followed by the confidence interval range in parentheses. RMSEA values can be elevated in cases of low degrees of freedom (Kenny et al., 2015).
*$P < .001$.

TABLE A.5
Correlation of latent variables.

Variable	Student engagement	Beliefs about CR instruction	Confidence in CR instruction	CR instruction	Ambitious instruction
Student engagement					
Beliefs about CR instruction	−.09				
Confidence in CR instruction	.11	.49			
CR instruction	.13	.34	.51		
Ambitious instruction	.32	.06	.3	.62	
Traditional instruction	.11	−.07	.12	.36	.65

TABLE A.6
District sample size and latent variable means.

Sample size by district	PL policy attri-butes	Curriculum policy attributes	Confidence in CR instruction	Beliefs about CR instruction	Ambitious instruc-tion	CR instruc-tion	Traditional instruction	Student engage-ment
35	3.4	2.6	5.7	3.7	1.4	1.0	1.6	1.6
13	3.4	2.8	8.4	4.8	1.2	1.6	1.7	1.6
10	3.7	3	7.7	3.9	2.2	2.2	2.4	1.9
29	3.1	2.8	6.7	3.7	1.6	1.4	2.1	1.5
30	3	2.6	7.3	3.7	1.5	1.4	2.1	1.5
25	3.8	3.1	7.0	2.9	2.0	1.6	2.5	1.9
9	4.3	3.3	7.1	3.8	1.6	1.5	2.1	1.7
10	3.5	2.7	6.8	2.9	1.5	1.2	2.1	1.3
86	3.1	2.7	6.0	3.0	1.5	1.0	2.1	1.8
142	3.5	3	7.4	3.5	1.6	1.4	2.1	1.6
37	3.5	3	7.3	3.7	1.3	1.1	1.6	1.9

PL, professional learning; CR, culturally responsive

Version of Record at: https://doi.org/10.1177/23328584241310429

Declaration of Conflicting Interests

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

Funding

This work was supported by the Bill & Melinda Gates Foundation (ID INV-000745).

Open Practices

The source materials for this article can be found at https://www.openicpsr.org/openicpsr/project/212322/version/V1/view

ORCID iDs

Laura M. Desimone https://orcid.org/0000-0003-0184-8997

Nick Bell https://orcid.org/0000-0001-9276-1883

Kirsten L. Hill https://orcid.org/0000-0001-9260-5178

Latrice Marianno https://orcid.org/0000-0001-6370-816X

References

Abdulrahim, N. A., & Orosco, M. J. (2020). Culturally responsive mathematics teaching: A research synthesis. *The Urban Review*, *52*(1), 1–25.

Agodini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness*, *3*(3), 199–253. https://doi.org/10.1080/19345741003770693

Alozie, N. M., Moje, E. B., & Krajcik, J. S. (2010). An analysis of the supports and constraints for scientific discussion in high school project-based science. *Science Education*, *94*(3), 395–427. https://doi.org/10.1002/sce.20365

Appleton, K. (1995). Student teachers' confidence to teach science: Is more science knowledge necessary to improve self-confidence? *International Journal of Science Education*, *17*(3), 357–369. https://doi.org/10.1080/0950069950170307

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. https://doi.org/10.3102/0034654315582066

Asparouhov, T., & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(1), 119–142.

Babinski, L. M., Amendum, S. J., Knotek, S. E., Sanchez, M., & Malone, P. (2018). Improving young English learners' language and literacy through teacher professional development: A randomized controlled trial. *American Educational Research Journal*, *55*(1), 117–143. https://doi.org/10.3102/0002831217732335

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher*, *25*(9), 6–14. https://doi.org/10.2307/1177151

Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a decade*. Rand Corporation.

Blazar, D. (2021). *Teachers of color, culturally responsive teaching, and student outcomes: Experimental evidence from the random assignment of teachers to classes* (EdWorkingPaper No. 21-501). Annenberg Institute at Brown University. https://doi.org/10.26300/jym0-wz02

Bollen, K. A. (1989). *Structural Equations With Latent Variables* (Vol. 210). John Wiley & Sons.

Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*. Sage.

Bollen, K. A., & Pearl, J. (2013) Eight myths about causality and structural equation models. In S. Morgan (Ed.), *Handbook of causal analysis for social research*. Springer.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15. https://doi.org/10.3102/0013189x033008003

Bottiani, J. H., Larson, K. E., Debnam, K. J., Bischoff, C. M., & Bradshaw, C. P. (2018). Promoting educators' use of culturally responsive practices: A systematic review of in-service interventions. *Journal of Teacher Education*, *69*(4), 367–385. https://doi.org/10.1177/0022487117722553

Bray-Clark, N., & Bates, R. (2003). Self-efficacy beliefs and teacher effectiveness: Implications for professional development. *Professional Educator*, *26*(1), 13–22. https://files.eric.ed.gov/fulltext/EJ842387.pdf

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. University of Chicago Press.

Burstein, L. (1995). *Validating national curriculum indicators*. Rand Corporation.

Byrd, C. M. (2016). Does culturally relevant teaching work? An examination from student perspectives. *Sage Open*, *6*(3), 2158244016660744. https://doi.org/10.1177/2158244016660744

Cherfas, L., Casciano, R., & Wiggins, M. A. (2021). It's bigger than hip-hop: Estimating the impact of a culturally responsive classroom intervention on student outcomes. *Urban Education*, *56*(10), 1748–1781. https://doi.org/10.1177/0042085918789729

Choppin, J., Roth McDuffie, A., Drake, C., & Davis, J. (2020). The role of instructional materials in the relationship between the official curriculum and the enacted curriculum. *Mathematical Thinking and Learning*, *24*(2), 123–148. https://doi.org/10.1080/10986065.2020.1855376

Civitillo, S., Juang, L. P., Badra, M., & Schachner, M. K. (2019). The interplay between culturally responsive teaching, cultural diversity beliefs, and self-reflection: A multiple case study. *Teaching and Teacher Education*, *77*, 341–351. https://doi.org/10.1016/j.tate.2018.11.002

Cliff, N. (1983). Cautions concerning the application of causal modelling methods. *Multivariate Behavioral Research*, *18*, 115–126. https://doi.org/10.1207/s15327906mbr1801_7

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, *12*, 311–329. https://doi.org/10.3102/01623737012003311

Comstock, M., Edgerton, A., & Desimone, L. M. (2022). Connecting policy to practice: How state and local policy environments relate to teachers' instruction. *Teachers College Record*, *124*(11), 82–116. https://doi.org/10.1177/01614681221143548

Comstock, M., Litke, E., Hill, K., & Desimone, L. M. (2023). A culturally responsive disposition: How professional

Version of Record at: https://doi.org/10.1177/23328584241310429

learning and teachers' beliefs about and self-efficacy for culturally responsive teaching relate to instruction. *AERA Open*, *9*. https://doi.org/10.1177/23328584221140092

Comstock, M., Kenneth, A. S., Camila, P., Erica, L., Kirsten, L. H., & Laura M. D. (2022). Typologizing Teacher Practice: How Teachers Integrate Culturally Responsive, Ambitious, and Traditional Teaching Approaches. (EdWorkingPaper: 22 -506). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/tnqf-qb36

Cooper, K. S. (2014). Eliciting engagement in the high school classroom: A mixed-methods examination of teaching practices. *American Educational Research Journal*, *51*(2), 363–402. https://doi.org/10.3102/0002831213507973

Copenhaver, J. F. (2001). Listening to their voices connect literary and cultural understandings: Responses to small group read-alouds of "Malcolm X: A Fire Burning Brightly." *New Advocate*, *14*(4), 343–359.

Covay, E., Desimone, L., Phillips, K., & Spencer, K. (2015). A new look at the opportunity to learn gap across race and income. *American Journal of Education*, *12*(2), 241–269. https://doi.org/10.1086/679392

Cruz, R. A., Manchanda, S., Firestone, A. R., & Rodl, J. E. (2020). An examination of teachers' culturally responsive teaching self-efficacy. *Teacher Education and Special Education*, *43*(3), 197–214. https://doi.org/10.1177/0888406419875194

Davis, E. A., & Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3–14. https://doi.org/10.3102/0013189x034003003

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267–297. https://doi.org/10.1002/pam.21818

Delpit, L. (2006). *Other people's children: Cultural conflict in the classroom*. New Press.

Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educational Psychological Measurement*, *77*(2), 185–203. https://doi.org/10.1177/0013164416658325

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Desimone, L. (2006). Consider the source: Response differences among teachers, principals and districts on survey questions about their education policy environment. *Educational Policy*, *20*(4), 640–676. https://doi.org/10.1177/0895904805284056

Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, *72*(3), 433–479. https://doi.org/10.3102/00346543072003433

Desimone, L., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society, & Education*, *7*(3), 252–263. https://doi.org/10.25115/psye.v7i3.515

Desimone, L., & Hill, K. (2017). Inside the Black box: Examining mediators and moderators of a middle school science intervention. *Educational Evaluation and Policy Analysis*, *39*(3), 511–536. https://doi.org/10.3102/0162373717697842

Desimone, L., & LeFloch, K. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, *26*(1), 1–22. https://doi.org/10.3102/01623737026001001

Desimone, L., & Long, D. (2010). Does conceptual instruction and time spent on mathematics decrease the student achievement gap in early elementary school? Findings from the Early Childhood Longitudinal Study (ECLS). *Teachers College Record*, *112*(12), 3024–3073. https://doi.org/10.1177/016146811011201206

Desimone, L., Smith, T., & Frisvold, D. (2010). Survey measures of classroom instruction: Comparing student and teacher reports. *Educational Policy*, *24*(2), 267–329. https://doi.org/10.1177/0895904808330173

Desimone, L., Wolford, T., & Hill, K. L. (2016). Research-practice: A practical conceptual framework. *AERA Open*, *2*(4). https://doi.org/10.1177/2332858416679599

Desimone, L., Smith, T. S., Hayes, S., & Frisvold, D. (2005). Beyond accountability and average math scores: Relating multiple state education policy attributes to changes in student achievement in procedural knowledge, conceptual understanding and problem solving in mathematics. *Educational Measurement: Issues and Practice*, *24*(4), 5–18. https://doi.org/10.1111/j.1745-3992.2005.00019.x

Desimone, L., Stornaiuolo, A., Flores, N., Pak, K., Edgerton, A. K., Nichols, T. P., Plummer, E., & Porter, A. (2019). Successes and challenges of the "new" college- and career-ready standards: Seven implementation trends. *Educational Researcher*, *48*(3), 167–178. https://doi.org/10.3102/0013189x19837239

Edgerton, A. K., & Desimone, L. M. (2018). Teacher implementation of college- and career-readiness standards: Links among policy, instruction, challenges, and resources. *AERA Open*, *4*(5), 1–22. https://doi.org/10.1177/2332858418806863

Edgerton, A. K., Desimone, L. M., & Harkin, J. (2020). *Appropriate and useful: Principal attitudes towards K–12 standards and policies*. Teachers College Record. https://journals.sagepub.com/pb-assets/cmscontent/TCZ/Research%20Notes%20Collection/2020%20Research%20Notes/Appropriate%20and%20Useful-%20Principal%20Attitudes%20Toward%20K%E2%80%9312%20Standards%20and%20Policies%20-1656010548.pdf

Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods*, *8*(3), 322.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*(3), 430–457.

Duke, N. K., Halvorsen, A-L., Strachan, S. L., Kim, J., & Konstantopoulos, S. (2021). Putting PjBL to the test: The impact of project-based learning on second graders' social studies and literacy learning and motivation in low-SES school settings. *American Educational Research Journal*, *58*(1), 160–200. https://doi.org/10.3102/0002831220929638

Feger, M-V. (2006). "I want to read": How culturally relevant texts increase student engagement in reading. *Multicultural Education*, *13*(3), 18–19. https://link.gale.com/apps/doc/A143629436/AONE?u=anon~e71bf9a0&sid=googleScholar&xid=e07d43f1

Version of Record at: https://doi.org/10.1177/23328584241310429

Fitchett, P. G., Starker, T. V., & Salyers, B. (2012). Examining culturally responsive teaching self-efficacy in a preservice social studies education course. *Urban Education*, *47*(3), 585–611. https://doi.org/10.1177/0042085912436568

Fowler Jr, J. F. (1995). *Improving survey questions: Design and evaluation*. Sage.

Freeman, D. T., & Porter, A. C. (1989). Do textbooks dictate the content of mathematics instruction in elementary schools? *American Educational Research Journal*, *26*(3), 403–421. https://doi.org/10.3102/00028312026003403

Freire, P. (2000). *Pedagogy of freedom: Ethics, democracy, and civic courage*. Rowman & Littlefield.

Garet, M. S., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Analysis of a national sample of teachers. *American Education Research Journal*, *38*(3), 915–945. https://doi.org/10.3102/00028312038004915

Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H., Doolittle, F., Zhu, P., Sztejnberg, L., & Silverberg, M. (2008). *The impact of two professional development interventions on early reading instruction and achievement* (NCEE 2008-4030). US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee/pdf/20084030.pdf

Garet, M. S., Heppen, J. B., Walters, K., Parkinson, J., Smith, T. M., Song, M., Garrett, R., Yang, R., Borman, G. D., & Wei, T. E. (2016). *Focusing on mathematical knowledge: The impact of content-intensive teacher professional development* (NCEE 2016-4010). US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. http://ies.ed.gov/ncee/pubs/20164010/pdf/20164010.pdf

Gay, G. (2021). Culturally responsive teaching: Ideas, actions, and effects. In R. Milner, & K. Lomotey (Eds.), *Handbook of urban education* (2nd ed.) (pp. 212–233). Routledge. https://doi.org/10.4324/9780429331435

Gay, G. (2018). *Culturally responsive teaching: Theory, research, and practice* (3rd ed.). Teachers College Press.

Gay, G. (2010). Acting on beliefs in teacher education for cultural diversity. *Journal of Teacher Education*, *61*(1–2), 143–152. https://doi.org/10.1177/0022487109347320

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education*, *52*(2), 106–116. https://doi.org/10.1177/0022487102053002003

Gay, G. (2000). *Culturally responsive teaching: Theory, research, & practice*. Teachers College Press.

Gerbing, D. W., & Anderson, J. C. (1988). The effect of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis. *Multivariate Behavioral Research*, *20*, 255–271. https://doi.org/10.1207/s15327906mbr2003_2

Ginsberg, M. B. (2015). *Excited to learn: Motivation and culturally responsive teaching*. Corwin.

Goldin, S., O'Neill, M. K., Naik, S. S., & Zaccarelli, F. G. (2019). Supporting students' learning practices: Redefining participation and engagement. *Elementary School Journal*, *119*(3), 417–442. https://doi.org/10.1086/701654

Gomer, B., Jiang, G., & Yuan, K. H. (2018). New effect size measures for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(3), 371–389. https://doi.org/10.1080/10705511.2018.1545231

Greenleaf, C. L., Litman, C., Hanson, T. L., Rosen, R., Boscardin, C. K., Herman, J., & Schneider, S. A. (2011). Integrating literacy and science in biology: Teaching and learning impacts of reading apprenticeship professional development. *American Educational Research Journal*, *48*(3), 647–717. https://doi.org/10.3102/0002831210384839

Griner, A. C., & Stewart, M. L. (2013). Addressing the achievement gap and disproportionality through the use of culturally responsive teaching practices. *Urban Education*, *48*(4), 585–621. https://doi.org/10.1177/0042085912456847

Guerra, P. L., & Wubbena, Z. C. (2017). Teacher beliefs and classroom practices cognitive dissonance in high stakes test-influenced environments. *Issues in Teacher Education*, *26*(1), 35–51.

Gutstein, E. (2003). Teaching and learning mathematics for social justice in an urban, Latino school. *Journal for Research in Mathematics Education*, *34*(1), 37–73.

Gutstein, E., Lipman, P., Hernandez, P., & De Los Reyes, R. (1997). Culturally relevant mathematics teaching in a Mexican American context. *Journal for Research in Mathematics Education*, *28*(6), 709–737. https://doi.org/10.2307/749639

Guyton, E. M., & Wesche, M. V. (2005). The multicultural efficacy scale: Development, item selection, and reliability. *Multicultural Perspectives*, *7*(4), 21–29. https://doi.org/10.1207/s15327892mcp0704_4

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (pls-sem) using R*. Springer.

Harbour, K. E., Evanovich, L. L., Sweigart, C. A., & Hughes, L. E. (2015). A brief review of effective teaching practices that maximize student engagement. *Preventing School Failure: Alternative Education for Children and Youth*, *59*(1), 5–13. https://doi.org/10.1080/1045988x.2014.919136

Heckman, J. J. (2005). The scientific model of causality. *Sociological Methodology*, *35*(1), 1–97.

Hill, H. C., Litke, E., & Lynch, K. (2018). Learning lessons from instruction: Descriptive results from an observational study of urban elementary classrooms. *Teachers College Record*, *121*(12), 1–21. https://doi.org/10.1177/016146811812001207

Hill, K. H., Desimone, L., Wolford, T., Reitano, A., & Porter, A. (2023). Inside school turnaround: What drives success? *Journal of Educational Change*, *24*(3), 393–423. https://doi.org/10.1007/s10833-022-09450-w

Hill, M. L. (2009). Wounded healing: Forming a storytelling community in hip-hop lit. *Teachers College Record*, *111*, 248–293. https://doi.org/10.1177/016146810911100109

Hodge, E. M., & Stosich, E. L. (2022). Accountability, alignment, and coherence: How educators made sense of complex policy environments in the Common Core Era. *Educational Evaluation and Policy Analysis*, *44*(4), 543–566. https://doi.org/10.3102/01623737221079650

Horn, I., & Garner, B. (2022). *Teacher learning of ambitious and equitable mathematics instruction: A sociocultural approach*. Routledge.

Version of Record at: https://doi.org/10.1177/23328584241310429

Howard, T. C., & Rodriguez-Minkoff, A. C. (2017). Culturally relevant pedagogy 20 years later: Progress or pontificating? What have we learned, and where do we go?. *Teachers College Record*, *119*(1), 1–32.

Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). Routledge.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hubert, T. L. (2014). Learners of mathematics: High school students' perspectives of culturally relevant mathematics pedagogy. *Journal of African American Studies*, *18*(3), 324–336. https://doi.org/10.1007/s12111-013-9273-2

International Literacy Association. (2019). *Children's rights to excellent literacy instruction [Position statement]*. Newark, DE: Author.

Irvine, J. (2020). Positively influencing student engagement and attitude in mathematics through an instructional intervention using reform mathematics principles. *Journal of Education and Learning*, *9*(2), 48–75. https://doi.org/10.5539/jel.v9n2p48

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Kenny, D., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling*, *10*(3), 333–351. https://doi.org/10.1207/S15328007SEM1003_1

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507.

Kline, R. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.

Kraft, M. A., & Hill, H. C. (2020). Developing ambitious mathematics instruction through web-based coaching: A randomized field trial. *American Educational Research Journal*, *57*(6), 2378–2414. https://doi.org/10.3102/0002831220916840

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*, 465–491. https://doi.org/10.3102/00028312032003465

Ladson-Billings, G. (2009). *The dream-keepers: Successful teachers of African American children* (2nd ed.). Jossey-Bass.

Lai, K., & Green, S. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate Behavioral Research*, *51*(2–3), 220–239. https://doi.org/10.1080/00273171.2015.1134306

Lampert, M. (1992). Practices and problems in teaching authentic mathematics. In F. Oser, A. Dick, & J. L. Patry (Eds.), *Effective and responsible teaching: The new synthesis* (pp. 295–314). Jossey-Bass.

Langlie, M. L. (2008). *The effect of culturally relevant pedagogy on the mathematics achievement of Black and Hispanic high school students* [Doctoral dissertation]. Proquest Dissertations and Theses Database (UMI No. 3304098).

Larson, K. E., & Bradshaw, C. P. (2017). Cultural competence and social desirability among practitioners: A systematic review of the literature. *Children and Youth Services Review*, *76*, 100–111. https://doi.org/10.1016/j.childyouth.2017.02.034

Le, V-N., Lockwood, J. R., Stecher, B. M., Hamilton, L. S., & Martinez, J. F. (2009). A longitudinal investigation of the relationship between teachers' self-reports of reform-oriented instruction and mathematics and science achievement. *Educational Evaluation and Policy Analysis*, *31*(3), 200–220. https://doi.org/10.3102/0162373709336238

Lefstein, A. (2008). Changing classroom practice through the English National Literacy Strategy: A micro-interactional perspective. *American Educational Research Journal*, *45*(3), 701–737. https://doi.org/10.3102/0002831208316256

Litke, E. (2020). The nature and quality of algebra instruction: Using a content-focused observation tool as a lens for understanding and improving instructional practice, cognition, and instruction. *Cognition and Instruction*, *38*(1), 57–86. https://doi.org/10.1080/07370008.2019.1616740

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis*, *41*(3), 260–293. https://doi.org/10.3102/0162373719849044

MaCallum, R. C. (2003). 2001 Presidential address: Working with imperfect models. *Multivariate Behavioral Research*, *38*, 113–139. https://doi.org/10.1207/S15327906MBR3801_5

Markowitz, A. J. (2018). Changes in school engagement as a function of no Child Left Behind: A comparative interrupted time series analysis. *American Educational Research Journal*, *55*(4), 721–760. https://doi.org/10.3102/0002831218755668

Marks, H. M. (2000) Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37*(1), 153–184. https://doi.org/10.3102/00028312037001153

Marsh, J. A., Kerr, K. A., Ikemoto, G. S., Darilek, H., Suttorp, M., Zimmer, R. W., & Barney, H. (2005). *The role of districts in fostering instructional improvement lessons from three urban districts partnered with the institute for learning*. Rand Corporation.

Masters, J., De Kramer, R. M., O'Dwyer, L. M., Dash, S., & Russell, M. (2010). The effects of online professional development on fourth grade English language arts teachers' knowledge and instructional practices. *Journal of Educational Computing Research*, *43*(3), 355–375. https://doi.org/10.2190/ec.43.3.e

Mayer, D. P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, *21*(1), 29–45. https://doi.org/10.2307/1164545

McDonald, R. P. (2004). The specific analysis of structural equation models. *Multivariate Behavioral Research*, *39*(4), 687–713.

McDonald, R., & Ho, M. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, *7*(1), 64–82. https://doi.org/10.1037/1082-989X.7.1.64

Meehl, P. E., & Waller, N. G. (2002). The path analysis controversy: a new statistical approach to strong appraisal of verisimilitude. *Psychological Methods*, *7*(3), 283.

Morgan, S. (2013). *Handbook of causal analysis for social research (Handbooks of sociology and social research)*. Springer.

Muthén Bengt, O. (2002). Beyond sem: General latent variable modeling. *Behaviormetrika*, *29*(1), 81–117. https://doi.org/10.2333/bhmk.29.81

Version of Record at: https://doi.org/10.1177/23328584241310429

National Center for Education Statistics. (2019). *Characteristics of public-school teachers*. US Department of Education, Digest of Education Statistics. https://nces.ed.gov/programs/coe/indicator/clr

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author. No Child Left Behind Act, 20 U.S.C. §6301 (2001). https://www.nctm.org/Standards-and-Positions/Principles-and-Standards/

National Mathematics Advisory Panel. (2008). *The final report of the national mathematics*. https://files.eric.ed.gov/fulltext/ED500486.pdf

Nichols, P., Desimone, L., & Edgerton, A. (2021). Smart power in standards implementation after No Child Left Behind. *American Journal of Education*, *128*(1), 147–169. https://doi.org/10.1086/716463

Nykiel-Herbert, B. (2010). Iraqi refugee students: From a collection of aliens to a community of learners--the role of cultural factors in the acquisition of literacy by iraqi refugee students with interrupted formal education. *Multicultural Education*, *17*(3), 2–14.

Opfer, D. V., Kaufman, J. H., & Thompson, L. E. (2021). *Implementation of K–12 state standards for mathematics and English language arts and literacy: Findings from the American teacher panel* (RR-1529-1-HCT). Rand Corporation.

Osborne, J. F., Borko, H., Fishman, E., Gomez Zaccarelli, F., Berson, E., Busch, K. C., Reigh, E., & Tseng, A. (2019). Impacts of a practice-based professional development program on elementary teachers' facilitation of and student engagement with scientific argumentation. *American Educational Research Journal*, *56*(4), 1067–1112. https://doi.org/10.3102/0002831218812059

Pak, K., Desimone, L., & Garcia Saldivar, E. (2020). The adaptive challenges of curriculum implementation: Insights for educational leaders driving education reform. *AERA Open*, *6*(2). https://doi.org/10.1177/2332858420932828

Pak, K., McLaughlin, J., Saldivar Zimmerman, E., & Desimone, L. M. (2021). "Boots on the ground": The authority-power dynamic of regional service centers in the standards era. *Journal of Educational Policy*, *29*, 31. https://doi.org/10.14507/epaa.29.5289

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. https://doi.org/10.3102/0034654319840359

Pellegrini, M., Lake, C., Neitzel, A., & Slavin, R. E. (2021). Effective programs in elementary mathematics: A meta-analysis. *AERA Open*, *7*(1), 1–29. https://doi.org/10.1177/2332858420986211

Penuel, W. R., Gallagher, L. P., & Moorthy, S. (2011). Preparing teachers to design sequences of instruction in earth systems science: A comparison of three professional development programs. *American Educational Research Journal*, *48*(4), 996–1025. https://doi.org/10.3102/0002831211410864

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal*, *44*(4), 921–958. https://doi.org/10.3102/0002831207308221

Phuntsog, N. (2001). Culturally responsive teaching: What do selected United States elementary school teachers think? *Intercultural Education*, *12*(1), 51–64. https://doi.org/10.1080/14675980125627

Polikoff, M. (2021). *Beyond standards: The fragmentation of education governance and the promise of curriculum reform*. Harvard Education Press.

Polikoff, M. S. (2015). How well aligned are textbooks to the Common Core Standards in mathematics? *American Educational Research Journal*, *52*(6), 1185–1211. https://doi.org/10.3102/0002831215584435

Porter, A. C. (1994). National standards and school improvement in the 1990s: Issues and promise. *American Journal of Education*, *102*, 421–449. https://doi.org/10.1086/444081

Porter, A. C., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1986). *Research series no. 179: Content determinants*. Institute for Research on Teaching, Michigan State University. https://edwp.educ.msu.edu/research/wp-content/uploads/sites/10/2020/11/rs179.pdf

Portes, P. R., González Canché, M., Boada, D., & Whatley, M. (2018). Early evaluation findings from the instructional conversation study: Culturally responsive teaching outcomes for diverse learners in elementary school. *American Educational Research Journal*, *55*(3), 488–531. https://doi.org/10.3102/0002831217741089

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. https://doi.org/10.1177/0143034307078090

Raykov, T. (2011). Evaluation of convergent and discriminant validity with multitrait–multimethod correlations. *British Journal of Mathematical and Statistical Psychology*, *64*(1), 38–52. https://doi.org/10.1348/000711009X478616

Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientations toward mathematics curriculum materials: Implications for teacher learning. *Journal for Research in Mathematics Education*, *35*(5), 352–388. https://doi.org/10.2307/30034820

Remillard, J. T., Harris, B., & Agodini, R. (2014). The influence of curriculum material design on opportunities for student learning. *ZDM*, *46*(5), 735–749. https://doi.org/10.1007/s11858-014-0585-z

Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. P. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal*, *47*(4), 833–878. https://doi.org/10.3102/0002831210367426

Roth, K. J., Wilson, C. D., Taylor, J. A., Stuhlsatz, M. A. M., & Hyidsten, C. (2019). Comparing the effects of analysis-of-practice and content-based professional development on teacher and student outcomes in science. *American Educational Research Journal*, *56*(4), 1217–1253. https://doi.org/10.3102/0002831218814759

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. *Teachers College Record*, *104*(8), 1525–1567.

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. A*sia-Pacific Journal of Teacher Education*, *39*(3), 183–198. https://doi.org/10.1080/1359866X.2011.588311

Version of Record at: https://doi.org/10.1177/23328584241310429

Seriki, V. D., & Brown, C. T. (2017). A dream deferred: A retrospective view of culturally relevant pedagogy. *Teachers College Record*, *119*(1), 1–8. https://doi.org/10.1177/016146811711900101

Siwatu, K. O. (2007). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education*, *23*(7), 1086–1101. https://doi.org/10.1016/j.tate.2006.07.011

Sleeter, C. E. (2012). Confronting the marginalization of culturally responsive pedagogy. *Urban education*, *47*(3), 562–584.

Smith, T. M., Desimone, L. M., & Ueno, K. (2005). "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school math. *Educational Evaluation and Policy Analysis*, *27*(1), 75–109. https://doi.org/10.3102/01623737027001075

Smith, T. M., Desimone, L. M., Zeidner, T., Dunn, A. C., Bhatt, M., & Rumyantseva, N. L. (2007). Inquiry-oriented instruction in science: Who teaches that way? *Educational Evaluation and Policy Analysis*, *29*(3), 169–199. https://doi.org/10.3102/0162373707306025

Souryasack, R., & Lee, J. S. (2007). Drawing on students' experiences, cultures, and languages to develop English language writing: Perspectives from three Lao heritage middle school students. *Heritage Language Journal*, *5*(1), 79–97. https://doi.org/10.46538/hlj.5.1.4

Spillane, J. P., & Jennings, N. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record*, *98*(3), 449–481. https://doi.org/10.1177/016146819709800303

Spillane, J. P., Hopkins, M., & Sweet, T. M. (2018). School district educational infrastructure and change at scale: Teacher peer interactions and their beliefs about mathematics instruction *American Educational Research Journal*, *55*(3), 532–571. https://doi.org/10.3102/0002831217743928

Spillane, J. P., Blaushild, N. L., Neumerski, C. M., Seelig, J. L., & Peurach, D. J. (2022). Striving for coherence, struggling with incoherence: A comparative study of six educational systems organizing for instruction. *Educational Evaluation and Policy Analysis*, *44*(4), 567–592. https://doi.org/10.3102/01623737221093382

Stecher, B., & Borko, H. (2002). Integrating findings from surveys and case studies: Examples from a study of standards-based education reform. *Journal of Education Policy*, *17*(5), 547–69. https://doi.org/10.1080/02680930210158311

Stornaiuolo, A., Desimone, L., & Polikoff, M. (2023). "The good struggle" of flexible specificity: Districts balancing specific guidance with autonomy to support standards-based instruction. *American Education Research Journal*, *60*(3), 521–561. https://doi.org/10.3102/00028312231161037

Swackhamer, L. E., Koellner, K., Basile, C., & Kimbrough, D. (2009). Increasing the self-efficacy of in-service teachers through content knowledge. *Teacher Education Quarterly*, *36*(2), 63–78. https://www.jstor.org/stable/23479252

Tarka, P. (2018). An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences. *Quality & Quantity*, *52*, 313–354.

Taylor, J. A., Getty, S. R., Kowalski, S. M., Wilson, C. D., Carlson, J., & Van Scotter, P. (2015). An efficacy trial of research-based curriculum materials with curriculum-based professional development. *American Educational Research Journal*, *52*(5), 984–1017. https://doi.org/10.3102/0002831215585962

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*, *50*(1), 29–34. https://doi.org/10.1016/j.jaci.2012.05.050

Turner, J. C., Christensen, A., Kackar-Cam, H. Z., Trucano, M., & Fulmer, S. M. (2014). Enhancing students' engagement: Report of a 3-year intervention with middle school teachers. *American Educational Research Journal*, *51*(6), 1195–1226. https://doi.org/10.3102/0002831214532515

Wang, M.-T., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement, and academic achievement in middle school. *American Educational Research Journal*, *47*(3), 633–662. https://doi.org/10.3102/0002831209361209

Yang, R., Porter, A. C., Massey, C., Merlino, F. J., & Desimone, L. M. (2020). Curriculum-based teacher professional development in middle school science: A comparison of training focused on cognitive science principles versus content knowledge. *Journal of Research in Science Teaching*, *57*(4), 536–566. https://doi.org/10.1002/tea.21605

Yoo, W., Mayberry, R., Bae, S., Singh, K., Peter He, Q., & Lillard, J. W. Jr. (2014). A study of effects of multicollinearity in the multivariable analysis. *International Journal of Applied Science Technology*, *4*(5), 9–19. https://pmc.ncbi.nlm.nih.gov/articles/PMC4318006/

Yuan, K., Le, V. N., McCaffrey, D. F., Marsh, J. A., Hamilton, L. S., Stecher, B. M., & Springer, M. G. (2013). Incentive pay programs do not affect teacher motivation or reported practices: Results from three randomized studies. *Educational Evaluation and Policy Analysis*, *35*(1), 3–22. https://doi.org/10.3102/0162373712462625

Zee, M., & Koomen, H. M. Y. (2016). Teacher self-efficacy and iI effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, *86*(4), 981–1015. https://doi.org/10.3102/0034654315626801

Authors

LAURA M. DESIMONE is the L. Sandra and Bruce L. Hammonds Professor at the University of Delaware. She studies education policy implementation and how best to support leaders and teachers in improving instruction in ways that bolster student engagement and learning.

NICHOLAS S. BELL is an assistant professor in special education at the University at Albany. His research focuses on studying anti-racist education practices and policies and their impact in terms of identifying and shaping equitable experiences for students in general and special education.

ARIELLE LENTZ is a PhD student studying education statistics and research methods with a secondary specialization in sociocultural and community-based approaches at the University of Delaware. Her research focuses on evidence use, funding, and equity in out-of-school time.

Version of Record at: https://doi.org/10.1177/23328584241310429

KIRSTEN HILL is a consultant specializing in survey design and impact evaluation at KLH Consulting, located in Sarasota, FL. Her research focuses on amplifying community voices through accessible and actionable data, with an emphasis on creating meaningful impact across sectors such as education, philanthropy, and social change.

LATRICE MARIANNO is an assistant professor in the Department of Educational Leadership at Southern Illinois University–Edwardsville. Her research focuses on supporting educational leaders to center equity within their school improvement efforts through policy and professional learning.