Mind the Gaps: Differences in How Teachers, Principals, and Districts Experience College- and Career-Readiness Policies

ADAM KIRK EDGERTON
University of Pennsylvania

LAURA M. DESIMONE
University of Delaware

Critics of standards-based reform often cite an accountability policy environment that disproportionately affects teachers compared with principals and district officials. We directly examine this disproportionality. In our three study states of Texas, Ohio, and Kentucky, we use survey analysis to understand how policy environments for district officials, principals, and teachers differ. We find that in all three states, teachers report experiencing significantly more accountability than do principals. Teachers in every state also report significantly lower authority toward their state's standards. In Texas, these authority gaps predict less coverage of English language arts standards.

Standards-based reform in K–12 public education began in the 1960s and 1970s under a clear rationale: set standards, measure them, and hold people accountable for the results (Smith and O'Day 1991). In this early period, individual states deployed only minimum competency tests. Although standard setting has been ongoing over the nation's history (see Gamson 2015), standards-based reform clearly accelerated after the publication of *A Nation at Risk* (National Commission on Excellence in Education 1983). A narrative of decline in educational standards, school quality, and student achievement has fueled new standards-based reform legislation for several decades (Mehta 2013). Bureaucracies at the district, state, and federal levels have increasingly addressed issues

Electronically published June 20, 2019
American Journal of Education 125 (August 2019)

0195-6744/2019/12504-00XX$10.00

of low academic achievement through standard setting and measurement. The quantification of these achievement or opportunity gaps among different groups of students has justified various policies aimed at gap closure (Foucault and Sheridan 1975; Gutiérrez and Dixon-Román 2011). This study, rather than examine achievement gaps, uses a policy attributes theory to quantify policy perception gaps among teachers, principals, and district officials.[1] We relate these gaps to the instructional content emphasized by each state's K–12 standards to see whether these policies make it past the classroom door.

We provide a strong theoretical approach to make sense of competing views on the nature and demands of standards-based reform as well as an understanding of arguments both for and against these policies. The primary opponents of standards-based reform have been those advocating for more local and democratic control (e.g., Tampio 2016) as well as those concerned about excessive quantification and testing of students (e.g., Ladd 2017). Political groups from both the right and the left have at times brought up objections to the quantification of schooling (Mehta 2013) or to standardization that removes local control from city, town, and county governments (Cohen and Mehta 2017; Desimone et al. 2019). Proponents have looked to well-implemented standards as a means of ensuring equality of opportunity (Hodge 2018; Satz 2007).

Effective implementation, however, has been elusive. Policy makers rely on teachers and principals to understand their intent and make sense of the policy (Coburn and Talbert 2006; Spillane 2009), but teachers may resist top-down policies (Hargreaves 2010). Furthermore, educators who do not view their state's college- and career-readiness (CCR) standards as appropriate for special populations (e.g., English learners [ELs]) and students with disabilities (SWDs) may subvert implementation of new standards (Edgerton et al. forthcoming; Flores and Schissel 2014; Powell et al. 2013).[2] This study is among the first to attempt to quantify these complex relationships.

Although national surveys provide descriptions of teacher attitudes toward standards (Ingersoll et al. 2014) and opinion polling provides insights into long-term trends of broader public opinion (Henderson et al. 2017), little empirical evidence exists to compare and test teacher, principal, and district official attitudes. Many compelling qualitative studies demonstrate how teachers feel about the state of their profession (e.g., Dunn et al. 2017) but little quantitative survey

---

ADAM KIRK EDGERTON is a PhD candidate in education policy at the University of Pennsylvania Graduate School of Education. His research interests include the politics of standards-based reform, teacher policy, and collective bargaining agreements. LAURA M. DESIMONE is the director of research at the University of Delaware's College of Education and Human Development. She studies policy effects on teaching and learning.

work compares and contrasts teachers' experiences and views to those of higher-level administrators. These three groups—teachers, principals, and district officials—are often studied in isolation by looking at gaps between teachers and principals only or by looking at the relationships between schools and central offices (e.g., Berends et al. 2002). An exception is Desimone (2006), who compared teacher, principal, and district responses with standards-based reform across five states in the early 2000s, noting general agreement on perceptions about barriers to success with principals being significantly more positive about accountability than teachers.

One survey of teachers only, the RAND American Teacher Panel Survey, finds ongoing confusion around "too many grade-level topics as aligned with their standards, and they [teachers] cannot always identify the order in which standards are taught" (Kaufman, Opfer, et al. 2018, 27). Surveying teachers alone, however, does not allow us to examine whether principals and district administrators have vastly different perceptions of standards policy. Our study focuses on state representative samples rather than a national one and compares these groups within and across Texas, Ohio, and Kentucky. Our study design allows us to assess whether principals and district officials believe that they are providing sufficient guidance compared with teachers, and we can measure policy gaps among the three groups rather than a single group in isolation.

Previous qualitative and mixed methods research has triangulated the experiences of these three groups. Stosich (2016) reveals in case studies of three high-poverty urban schools that teachers need "direct support from school leaders and external partners" (1728). Hodge (2018) similarly studies standards implementation within a single district and finds that lower-tracked students do not receive the same standards-based curriculum as those on a higher track. Desimone (2006) analyzed interview data from all three groups' responses to an earlier iteration of standards-based reform, noting that respondents in all three groups attributed standards-based reform to increasing the focus on struggling learners, raising expectations for all students, and fostering more creative pedagogy. We know of no study, however, that uses survey data to examine how teachers, principals, and district officials simultaneously view and experience the current wave of standards-based reform.

Theoretical Framework

We use a robust theoretical framework known as the policy attributes theory (Porter 1994; Porter et al. 1988), which has been used for decades in education policy research to analyze systemic reform efforts (Clune 1993), comprehensive school reforms (Berends et al. 2002; Desimone 2002; Polikoff 2012), and research-practice partnerships (Desimone et al. 2016). The use of this theory

that allows us to quantify individual experience with standards-based policy is a major contribution of our study, as we provide a set of rigorously tested items for future research across additional states.

According to the policy attributes, successful policy implementation requires a policy environment that is specific, authoritative, consistent, powerful, and stable. *Specificity* refers to how extensive, detailed, or prescriptive a policy is (e.g., how much time a teacher should spend on content). *Authority* reflects how policies gain legitimacy and status through persuasion (e.g., rules or law, historical practice, charismatic leaders). *Consistency* captures the extent to which policies are aligned and how policies relate to and support each other (e.g., whether curricula, assessments, professional development, and evaluations are aligned with each other and the standards). *Power* reflects how policies are reinforced and enacted through systems of rewards and sanctions (e.g., dismissals for low performance on tests; pay for performance). *Stability* refers to the extent to which policies change or remain constant over time (e.g., how long standards and assessments will last). Related survey analysis using the attributes has revealed how authority is significantly predictive of standards-emphasized instruction (Edgerton and Desimone 2018). The attributes have also enabled us to understand how policies have shifted over time, particularly the ways in which district officials now avoid, at least rhetorically, test-based accountability (Desimone et al. 2019).

There are, of course, limitations to the use of the attributes. They strive to be objective and apolitical by analyzing the educational bureaucracy and the actors within it. The attributes implicitly position themselves as positive and do not consider complex political negotiations and principles of democratic governance (for why this might be a concern, see Tampio 2016). Practitioners also may not have an understanding of the interlocking network of policies that implicitly or explicitly influences their instruction. Another limitation is that the attributes neither specifically focus on race and equity nor do they illuminate how policies might disproportionately impact different groups of students. However, our specific focus on SWDs and ELs supplements our theoretical framework in this regard.

State Policy Environments, Trends, and Implications for Differences

Our survey data collection in the spring of 2016 occurred immediately following the passage of the Every Student Succeeds Act (ESSA) in December 2015 (Pub. L. No. 114–95 § 114 Stat. 1177), which increased the standards-based policy flexibility afforded by the federal government. Some states used this flexibility to either double down on or lessen test-based accountability (C-SAIL 2019). Our study states that were selected for their variation in size, geography,

and demographics, as well as their willingness to participate in a 5-year study, also represent a wide range of education policy environments. As shown in table 1, Kentucky, a more rural state with fewer than 1 million students and the highest poverty rate of the three states, is much smaller than Ohio, with 1.8 million students, and Texas, with 5 million students. Kentucky's and Ohio's K–12 students are overwhelmingly white, which is not the case in Texas. Finding no differences among the states would add to growing evidence that practitioner perceptions of standards-based policy have become nationalized through political rhetoric and social media (Edgerton and Desimone 2018; Polikoff et al. 2016; Supovitz 2017).

Briefly, we describe the most relevant policies in each state that could help explain the levels of the attributes: textbook policy (specificity and consistency), testing and evaluation (power), and standards legislation and implementation time lines (authority and stability). We compare states directly, mirroring our survey analysis. We do not discuss curricula, as none of the states in these studies

TABLE 1

Contextual Information and Policy Attributes for Kentucky, Ohio, and Texas

	Texas	Ohio	Kentucky
PreK–12 enrollments	4,934,366	1,754,191	690,634
School districts (*n*)	1,239	955	173
Public schools (*n*)	8,732	3,758	1,449
Public school teachers (*n*)	334,997	109,282	41,588
Unadjusted education spending per student ($)	8,788	11,224	10,426
% minority students	68.8	25.8	21.0
% students eligible for free or reduced-price lunch	50.1	42.5	60.0
% students with disabilities	9.0	14.8	13.3
% English language learners	15.0	2.1	3.6
% students in urban schools	42.3	18.9	21.5
% students in rural schools	15.6	22.2	35.3
Specificity	Medium	High	Medium
Consistency	High	High	High
Authority	High	Low	Medium
Power	Medium	High	Medium
Stability	Medium	Medium	High

NOTE.—*Specificity* refers to the specificity of the state's definition for college and career readiness (CCR). *Consistency* refers to whether high school students are required to take CCR-aligned courses to graduate. *Authority* refers to recent legislation supporting CCR standards. *Power* refers to CCR indicators in the accountability system (including SAT/ACT, AP tests, and industry-recognized credentials) as well as rewards and sanctions for low-performing schools. *Stability* refers to the extent to which a state has revised its assessments (C-SAIL 2018).

have a curriculum approval process (e.g., as Louisiana does; Kaufman, Cannon et al. 2018). We also strongly caution against tying these policies directly to our survey findings; they merely provide descriptive background.

Comparing State CCR Standards Policies

Texas fully adopted its CCR standards and associated tests starting in 2009–10, which preempted the rest of the country's shift toward the Common Core State Standards (CCSS). This lengthier time line compared with the other states may be contributing to significantly higher authority and stability (Edgerton and Desimone 2018). A state dominated by a single party with a history of independence from federal initiatives, Texas never adopted the CCSS and charted its own path through the Texas Essential Knowledge and Skills (TEKS). The TEKS standards were significantly more rigorous than the Ohio and Kentucky standards prior to their adoption of the CCSS (Carmichael et al. 2010). As for textbook adoption processes, this responsibility rests with individual school districts instead of a state approval board (Texas Education Agency 2019).

Ohio, in contrast, maintains a state approval process for textbooks, which relates to the attributes of specificity and consistency (Ohio Department of Education 2019). Ohio was also the last of these states to fully implement its CCR standards policies in 2013–14, and it has the strongest teacher accountability policies when comparing consequences for student performance (C-SAIL 2019). Ohio also allows for the least amount of local control, which we define as district-level decision-making power over textbooks and accountability.

Kentucky, with its high local control through its school-based decision making councils, contrasts sharply with Ohio (Kentucky Department of Education 2019). Kentucky fully implemented its standards-based policies in 2011–12 and saw major changes in the political climate surrounding standards-based reform. The state legislature and governor were far less supportive of standards-based reform in 2016 according to interviews (Desimone et al. 2019).

Using an interactive and comparative 50-state database of policies (C-SAIL 2018), we also classify in table 1 where we believe each state to fall in terms of the attributes. Based on prior survey analysis of teachers (Edgerton and Desimone 2018), our understanding of implementation trends using complimentary qualitative work (see Desimone et al. 2019), and this policy database, we consider Texas to have the tightest linkages between policy and instruction, followed by Kentucky, and then Ohio. Texas perhaps benefitted from not withdrawing from the Partnership for Assessment of Readiness for College and Careers as Ohio and Kentucky did (C-SAIL 2018).

In a related mixed methods study using both survey and interview data, we also identify implementation trends that are broadly applicable across all three

states. First, state and district officials in all three states described an increased emphasis on local control. This emphasis may be discouraging states from pursuing curricular adoption and alignment activities. Second, districts are consequently still struggling to adopt aligned curricula and resources. Building consistency and specificity requires sustained, coherent professional development (Desimone, 2002; Desimone et al. 2016). Third, standards are now accepted as a norm or institutional logic (DiMaggio and Powell 1983) in spite of social media backlash to the CCSS (Supovitz 2017). Finally, for ELs, national organizations such as the WIDA Consortium are providing guidance on standards-based instruction while tensions persist between standards and differentiation for SWDs (Desimone et al. 2019; Edgerton et al. forthcoming).

Hypotheses and Research Questions

We offer several hypotheses for relationships between policy gaps and instruction. We specifically focus on standards-emphasized instruction (explained in depth in the Methods section) as the primary measure of implementation and our outcome of interest. Across these states, it may be the case that as long as principals and districts demonstrate a strong belief (authority) in standards-based policies, teachers will teach more of the standards, even if they feel that policies are unfair or inappropriate (Honig 2008). But if a policy perception gap exists between teachers and administrators, this gap might negatively impact standards implementation (Spillane 2009).

Principals are key to any standards-based reform in ensuring aligned instruction through consistent messages and routines (Supovitz et al. 2009). To help teachers shift instruction to match new standards, "educational leaders need the skills to lead with data" (Goldring and Schuermann 2009, 12) as well as a deep understanding of the standards across subject areas (Stosich 2016). The principal can serve as a reliable connection in the notoriously disconnected profession of teaching. School cultures require a stable instructional core of teachers, and principals have an influence on whether teachers stay (Boyd et al. 2011) as well as their job satisfaction (Griffith 2004). In North Carolina, teacher ratings of the school environment depend on principals independent of other school and district contextual factors (Burkhauser 2017). Thus, when studying the efficacy of a policy environment on instruction, we would expect significant policy gaps between teachers and principals to impact schools negatively.

Policy gaps between teachers, principals, and districts imply a failure of communication, a lack of integration between instructional and administrative roles, a lack of agreement about whether the standards are appropriate for all students, or all of the above. We hypothesize that gaps show disproportionality in the policy system across these three groups of respondents and that these

gaps may lead to less standards-emphasized instruction. Identifying the gaps will illuminate areas of improvement for policy implementation, including clearer communication from the central office, better integration of professional development, and other opportunities to build greater consensus between administrators and practitioners.

We do not limit our examination of differences to teachers and principals, however. To address the persistent stereotype of the out-of-touch central office bureaucrat who treats schools uniformly rather than dynamically (Honig 2006, 2008, 2009), we also quantify and test the gaps between principals and districts as well as between teachers and districts. Administrative climate broadly speaking can determine whether novice teachers remain (Pogodzinski et al. 2012). Thus, finding disproportionality in the policy system by itself is an area of concern.

This article addresses the following research questions: (1) How are standards-based policy environments for teachers, principals, and district officials similar or different across and within states? (2) Do teachers and principals by state experience standards-based policy differently within their districts and within their schools? (3) Do teachers and district officials by state experience standards-based policy differently within their districts? Do principals and district officials by state experience standards-based policy differently within their districts? (4) To what extent do policy environment gaps between teachers and principals predict teachers' preparedness to teach standards and the extent to which they align their instruction to standards-emphasized content?

Study Design, Data, and Methods

We use representative surveys to quantify perceptions. Respondents may be reporting on the same items and ostensibly the same standards policies, but each has different "truths." Each group experiences policy in a different way, and we focus on these experiences—not the written policy itself but rather how each group interprets it. Survey self-reports on policy perceptions have been shown to be valid and reliable across multiple studies (Desimone 2002; Desimone and Le Floch 2004).

We conducted the survey during spring 2016. We identified districts using a stratified random sampling technique designed to ensure the sample was representative of districts in each state: 42 Texas districts, 42 Ohio districts, and 89 Kentucky districts. There were more Kentucky districts sampled because Westat administered the survey in partnership with the state department of education, and some questions were removed from Kentucky's teacher and principal surveys. District officials did not take the survey in that state. In contrast,

American Institutes for Research (AIR) administered the survey independently in Ohio and Texas and expended significant resources in boosting response rates.

In each district, we sampled up to two elementary schools and two high schools, ensuring representative samples of public, private, and charter schools based on state demographics. We did not include middle schools because of limited resources and sought to heighten the contrast between elementary schools and high schools in their experiences with standards policy. In each elementary school, we sampled two fifth-grade math teachers, two fourth-grade English language arts (ELA) teachers, one teacher of SWDs, and one teacher of English language learners (ELLs). We did not sample primary grades because fourth- and fifth-graders were most likely to take the state test in our study states. In each high school, we sampled two ELA teachers and one teacher in each of the following specialties or topics: SWDs, ELLs, algebra I, algebra II, and geometry. We chose these three math subjects because they are the most common high school math courses and because they enroll students likely to be required to take the state mathematics assessment.

In Ohio and Texas, respectively, 155 and 121 eligible districts were contacted. Among the eligible districts, 49 and 53, respectively, agreed to participate, but only 42 completed the survey in each state (conditional response rate: 85.7% and 79.2% in Ohio and Texas, respectively). In Kentucky, all 89 districts were included and surveyed. AIR contacted sample schools within the Ohio and Texas districts that agreed to participate. Of the 185 eligible principals (or designated staff) in Ohio, 111 responded (conditional response rate: 60.0%), 149 of 211 eligible principals in Texas responded (70.6%), and 179 of 354 eligible principals in Kentucky responded (50.6%). Of the eligible teachers, in Ohio, 417 of 654 sampled teachers responded (conditional response rate: 64.8%); in Texas, 603 of 1,089 (55.3%); in Kentucky, 740 of 1,890 (39.2%).

Survey Development, Scales, and Measures

We used multi-item composites for each measure to increase reliability and validity (Mayer 1999), while borrowing items on already validated national surveys where possible. When we needed to create or adapt items, we completed a cycle of development, expert review, cognitive interview, and review (see Desimone and Le Floch 2004).

We structured items with parallel wording across the district, principal, and teacher surveys. For example, we asked all three respondent groups, "Please indicate your agreement with the following statement about the [State CCR Standards]." We started with a baseline item to capture general education students with statements such as "[State CCR Standards] set appropriate expectations for student learning at each grade level." To contrast the general edu-

cator's response for students with and without disabilities, a later item stated "[State CCR Standards] set appropriate expectations for students with disabilities," such that identical wording of items was used to form multi-item composites comparing responses for students with and without disabilities. For the present analysis, we included the subset of items relevant to personnel perceptions of policy attributes related to standards implementation, teachers' focus on content emphasized and deemphasized in the transition from the standards of the No Child Left Behind (NCLB) (P.L. 107–110, 20 U.S.C. § 6319 (2002)) to current CCR standards, and teacher preparedness. We describe each set of items below.[3]

Measuring standards alignment.—To measure the alignment of standards to instruction as done in prior studies, we rely on the Surveys of Enacted Curriculum approach (Porter 2002). In this procedure, teachers reported on the topics and cognitive demands that they covered in class. We then mapped that content onto the topics and cognitive demands from each state's standards using trained content experts convened at AIR. Surveys that ask about particular behaviors can be reliable ways of measuring teachers' instruction (Desimone 2009; Mayer 1999). Furthermore, multiple studies have used this technique and found it to be valid and reliable (Blank 2002; Clune 1993; Polikoff et al. 2011; Rowan et al. 2004; Webb 2002, 2007).

Our goal was to create a measure of alignment with content emphasized in each state's standards. In other words, we wanted to understand if teachers were teaching more of the content emphasized by the standards and spending less time on content that their state's CCR standards did not emphasize. To generate lists of emphasized and deemphasized instructional content, we used the following process at AIR. Trained raters first coded the topics and cognitive demands of each state's NCLB-era content standards and compared those with topics and cognitive demands of that state's newest CCR standards. Then we identified the content with the greatest average proportional increases and decreases at each grade level. Items with the greatest increases were designated as "emphasized" content; items with the greatest decreases were designated as "deemphasized" content. We also expected that teachers experienced these instructional changes in the field rather than preservice, as the vast majority of respondents (91% in Texas and Ohio, and 88% in Kentucky) had more than 3 years of experience; the mean years of experience for respondents was 12 years (Texas and Kentucky) or 14 years (Ohio).

Policy attribute scales.—All of our independent variables consisted of averages of multiple items on a Likert scale of 1–4. To measure specificity, we asked respondents for their level of agreement with statements related to how detailed guidance was around standards implementation, including how much time teachers should spend on each content area and in what order. For authority, we asked respondents the extent of their agreement with statements that re-

flected their buy-in to their state's standards, such as if they thought the standards made learning relevant, if the standards were appropriate for their students, and if the standards gave the flexibility they needed to help students below grade level. For consistency, we asked the degree to which respondents believed curricula, assessments, professional development, evaluations, and other policies were aligned. For power, we asked about positive and negative repercussions for implementing the standards. For stability, we asked them to predict how long the standards and assessments would last in each state. Cronbach's alpha, which measures how well items correlate to form a composite (where 1 is the highest) ranged from .64 to .92; all values except for power (.64) were between .80 and .90. We attribute the lower alpha for power to the fact that the construct includes both rewards and sanctions—that is, districts may be more likely to have one or the other but not both.

Teachers' preparedness.—We asked math and ELA teachers, "How prepared did you feel to teach the standards to low-achieving students, students with individualized education plans (IEPs), and English language learners (ELLs)?" These three populations were the focus of our study, and we compared the response with low-achieving students more generally to see how general educators viewed their SWDs and ELLs. We asked teachers of SWDs and ELLs to report on their students only. For ELA and math teachers, we averaged across all three items to create a single preparedness score for each teacher respondent. The scale was 1–4 (1 = "I do not feel prepared"; 2 = "I feel slightly prepared"; 3 = "I feel moderately prepared"; 4 = "I feel well prepared").

Methods

We examined patterns of nonresponse in each state for areas of concern and found no significant differences among subject areas of teachers (math, ELA, SWD, and EL) responding to the survey. Throughout this article, we used the jackknife resampling procedure in Stata, so that all responses were weighted to be representative of the sample to correct for any nonresponse bias. We report robust standard errors obtained from this procedure throughout this article. Once we obtained state means, we used ANOVA tests to identify significant differences between our three states by respondent. We used a *t*-test of means where Kentucky data were not available.

After describing the average overall levels and seeing whether significant differences were meaningful based on effect sizes, we paired respondents by school and then by district to see whether averages were obscuring important differences happening at individual schools. We used a paired *t*-test for teachers and principals and then for teachers and district officials. Finally, we tested

to see whether gaps (absolute value) between teachers and principals could predict standards-emphasized instruction through an ordinary least squares (OLS) regression:

$$y = \alpha + \beta x_1 + \beta x_2 + \beta x_3 + \beta x_4 + \beta x_5,$$

where y represents standards-emphasized instruction, x_1 is the absolute value of the specificity difference between teachers and principals, x_2 is the absolute value of the consistency difference between teachers and principals, x_3 is the absolute value of the authority difference between teachers and principals, x_4 is the absolute value of the power difference between teachers and principals, and x_5 is the overall gap. Stability is omitted because stability questions were not included on the Kentucky principal surveys.

Results

We consider the overall levels of policy attributes as reported by teachers in each state, and then we turn to principals and district officials. These descriptive findings ground the interpretations for all of the findings that follow. Because all Likert scores on the scale range from 1 to 4, an unstandardized effect size of at least 0.50 is meaningful, as it shows movement from one category to the next, for example, somewhat disagree (2) to somewhat agree (3). We report the results in table 2.

Research Question 1

Teachers.—To return to our hypotheses, we expected to find significant differences both within states and across states because of differing state policy environments for textbooks. Teacher-reported specificity refers to the guidance and support provided for standards-based instruction. In terms of specificity, differences between the states are all statistically significant in table 2. Scores near 3.0 (somewhat agree) for both Texas (3.14) and Kentucky (2.75) teachers are substantively similar even if they are significantly different at the 0.05 level. But with a significantly lower score of 2.38 for Ohio with large effect sizes (ES = −0.37 compared with Kentucky; −0.76 compared with Texas), the Ohio teachers' reports of the specificity of their standards environment present one possible area of improvement. Ohio teachers may need more guidance (e.g., the specific order in which content should be taught) from administrators, and this finding matches our understanding of the state's policy environment (C-SAIL 2018).

TABLE 2

Average Attributes for Teachers, Principals, and Districts across All Three States

	n	TX	*n*	OH	*n*	KY	*F*	Sig.
Teachers:								
Specificity	585	3.14 (.07)	405	2.38 (.13)	436	2.75 (.04)	21.02**	All
Consistency	564	2.82 (.04)	379	2.71 (.04)	390	2.91 (.03)	5.87**	TX vs. OH, OH vs. KY
Authority	583	2.56 (.05)	402	2.30 (.04)	430	2.78 (.03)	26.59**	All
Power	586	2.68 (.07)	405	2.50 (.06)	428	2.56 (.03)	1.73	
Stability	579	2.51 (.07)	398	2.44 (.06)	400	2.29 (.06)	3.38*	TX vs. KY
Principals:								
Specificity	154	2.94 (.09)	110	2.60 (.12)	158	2.84 (.07)	3.28*	TX vs. OH
Consistency	154	2.83 (.06)	108	2.83 (.06)	166	2.52 (.05)	9.97**	TX vs. KY, OH vs. KY
Authority	153	2.95 (.07)	109	2.90 (.08)	178	3.18 (.06)	4.93*	TX vs. KY, OH vs. KY
Power	153	2.35 (.10)	110	2.38 (.12)	153	1.96 (.08)	5.86**	TX vs. KY, OH vs. KY
Stability	153	2.74 (.09)	109	2.83 (.11)	N/A	N/A	.64	
Districts:	*n*	TX	*n*	OH			*t*-Test	
Specificity	42	3.35 (.11)	42	3.03 (.18)			1.52	
Consistency	42	2.75 (.13)	42	2.74 (.08)			.07	
Authority	42	2.73 (.11)	42	2.50 (.13)			1.35	
Power	42	2.34 (.16)	42	1.96 (.11)			1.96	
Stability	42	2.92 (.17)	42	2.28 (.20)			2.44*	

NOTE.—Sig. = significant. TX = Texas; OH = Ohio; KY = Kentucky; N/A = not available.

* $p < .05$.

** $p < .01$.

Teachers in all three states report consistency scores between 2.71 and 2.91. Consistency measures the alignment of standards with curricula, assessments, and supportive resources. Ohio teachers report significantly less consistency (2.71 versus 2.82 for Texas and 2.91 for Kentucky), but the effect size was small (ES = −0.11 compared with Texas, −0.20 compared with Kentucky). Average consistency scores below 3.0 suggest teachers believe materials and assessments need to be considerably more aligned to the standards.

For authority—that is, the extent to which teachers believed in and thought the standards were appropriate and helpful for their students—all comparisons are statistically significant across states. Kentucky teachers report the highest authority (2.78), whereas Ohio teachers again have the lowest (2.30, ES = 0.48). There are no significant differences across states for power, and stability differences are only significant between Texas (2.51) and Kentucky (2.29) with a relatively small effect size (0.22).

Principals and district officials.—For principals, we find much more similarity between Texas and Ohio on every attribute, and only specificity is statistically significant between these two states (Texas: 2.94; Ohio: 2.60; ES = 0.34). Kentucky principals, however, report significantly less consistency (ES = −0.31) and power (ES = −0.39 compared with Texas, −0.43 compared with Ohio) than the other two states. Kentucky principals report significantly higher authority (3.18), but the effect size is relatively small (0.23 compared with Texas; 0.28 compared with Ohio). Stability questions were omitted from the Kentucky principal survey administration.

For districts, we use a *t*-test to compare Texas and Ohio, as we do not have district data for Kentucky. We find no significant differences except for stability, with Texas district officials rating their policy environments as significantly more stable than Ohio (2.92 versus 2.28). For the other attributes, specificity is the highest attribute (3.35 for Texas and 3.03 for Ohio) followed by consistency (2.75 for Texas and 2.74 for Ohio), authority (2.73 for Texas and 2.50 for Ohio), and power, which is lowest (2.35 for Texas and 1.96 for Ohio). Power questions ask about the rewards and sanctions district officials enacted for teachers and principals; district officials believe they punished or rewarded to a relatively small extent. As we will see, teachers do not hold similar views.

Research Question 2

As shown in table 3, in Texas, there are significant differences for authority and power only. Teachers report significantly lower authority (ES = −0.39) but significantly higher power (ES = 0.33) than principals. In Kentucky, this same pattern recurs—that is, teachers report significantly lower authority (ES = −0.40) and significantly higher power (ES = 0.60) than principals. In Ohio,

TABLE 3

Within-State Policy Environment Comparisons of Teachers, Principals, and Districts

	n T	Mean	*n* P	Mean	*n* D	Mean	*F*	Sig.
Texas:								
Specificity	585	3.14	158	2.94	41	3.35	1.54	
		(.07)		(.09)		(.11)		
Consistency	564	2.82	166	2.83	42	2.75	.13	
		(.04)		(.06)		(.13)		
Authority	583	2.56	178	2.95	42	2.73	8.20**	T vs. P
		(.05)		(.07)		(.11)		
Power	586	2.68	153	2.35	42	2.34	3.47*	T vs. P
		(.07)		(.10)		(.16)		
Stability	579	2.51	153	2.74	41	2.92	2.36	
		(.07)		(.09)		(.17)		
Ohio:								
Specificity	405	2.38	110	2.60	42	3.03	1.69	
		(.13)		(.12)		(.18)		
Consistency	379	2.71	108	2.83	42	2.74	1.13	
		(.04)		(.06)		(.08)		
Authority	402	2.30	109	2.90	42	2.50	23.58**	T vs. P,
		(.04)		(.08)		(.13)		T vs. D
Power	405	2.50	110	2.38	42	1.96	4.08*	T vs. D
		(.06)		(.12)		(.11)		
Stability	398	2.44	109	2.83	42	2.28	5.36**	T vs. P,
		(.06)		(.11)		(.20)		P vs. D.
	n T	Mean	*n* P	Mean			*t*-Test	
Kentucky:								
Specificity	436	2.75	158	2.84			1.14	
		(.04)		(.07)				
Consistency	390	2.91	166	2.52			6.92**	
		(.03)		(.05)				
Authority	430	2.78	159	3.18			6.52**	
		(.03)		(.06)				
Power	428	2.56	153	1.96			8.66**	
		(.03)		(.08)				
Stability	400	2.29	N/A	N/A				
		(.06)		N/A				

NOTE.—T = teachers; P = principals; D = districts; Sig. = significant; N/A = not available.

* $p < .05$.

** $p < .01$.

teachers report significantly lower authority than principals (ES = −0.6) and significantly higher power than districts (ES = 0.56).

In Ohio only, district officials experience significantly lower authority (ES = −0.55) compared with principals. In Kentucky only, principals experience significantly lower consistency than teachers (ES = −0.39).

As for cross-state patterns, it is notable that across all three of our states, teachers report significantly lower authority and significantly higher power than either principals (in Texas and Kentucky) or district administrators (in Ohio). These gaps transcend state lines and state policy differences.

Research Question 3

The previous analysis of state averages has the potential to obscure important variations in types of schools and districts within and across states. To discern these variations, we examine the distribution of these gaps between teachers and principals and teachers and districts, and we find them to be normally distributed both within and across states. Thus, we decided on a paired *t*-test, where teachers are averaged by school and then by district, and then we compare them with their own responding principal and district official. In table 4, we report standard deviations instead of standard errors.

For the power and authority cross-state findings, the results hold constant when aggregating teachers by school and comparing them against their own principals. A few new and significant findings do emerge when not using state averages. Texas teachers perceive significantly higher specificity than did their principals (ES = 0.17). Ohio teachers report significantly less stability than their principals (ES = −0.45), and Kentucky teachers see significantly more consistency than their principals (ES = 0.38). Most importantly, the pattern of power and authority gaps between teachers and principals remains durable when matching teachers to their principals.

When comparing teachers to their own district officials, significant power and authority gaps again recur (see table 4). Teachers in Texas and Ohio perceive significantly lower authority (ES = −0.19 for Texas and −0.34 for Ohio) and significantly higher power (ES = −0.24 for Texas and 0.36 for Ohio) than district officials. A new finding does emerge for specificity that did not appear when simply averaging: teachers report significantly less specificity than district officials (ES = −0.41 in Texas and −0.67 in Ohio). This finding means that districts believe they are providing significantly more specific guidance than teachers report experiencing. Texas teachers also report significantly less stability than district officials (ES = −0.46).

TABLE 4

Paired t-Test Comparisons of Teachers versus Principals by School, Teachers versus District Officials by District

	n	Teacher	Principal	Difference
Schools:				
Texas:				
Specificity	144	3.16	2.99	.17*
		(.60)	(.72)	
Consistency	142	2.88	2.79	.09
		(.43)	(.50)	
Authority	143	2.58	2.97	−.39**
		(.49)	(.48)	
Power	143	2.72	2.45	.27**
		(.46)	(.84)	
Stability	142	2.56	2.74	−.18
		(.67)	(.96)	
Ohio:				
Specificity	107	2.40	2.55	−.15
		(.71)	(.79)	
Consistency	105	2.74	2.80	−.06
		(.32)	(.54)	
Authority	107	2.35	2.87	−.52**
		(.37)	(.48)	
Power	107	2.53	2.39	.13
		(.43)	(.89)	
Stability	106	2.41	2.86	−.45**
		(.58)	(1.06)	
Kentucky:				
Specificity	104	2.80	2.90	−.10
		(.62)	(.62)	
Consistency	104	2.93	2.54	.38**
		(.44)	(.71)	
Authority	104	2.79	3.20	−.41**
		(.46)	(.69)	
Power	102	2.55	1.96	.59**
		(.54)	(.75)	
Districts:				
Texas:				
Specificity	41	2.96	3.37	−.41**
		(.52)	(.66)	
Consistency	41	2.83	2.82	.01
		(.30)	(.52)	
Authority	41	2.54	2.73	−.19*
		(.30)	(.58)	
Power	41	2.69	2.45	.24*
		(.31)	(.69)	
Stability	40	2.52	2.99	−.46**
		(.56)	(.84)	

TABLE 4 (*Continued*)

	n	Teacher	Principal	Difference
Ohio:				
Specificity	42	2.20	2.86	−.67**
		(.50)	(.75)	
Consistency	42	2.75	2.85	−.10
		(.28)	(.57)	
Authority	42	2.28	2.62	−.34**
		(.29)	(.59)	
Power	42	2.44	2.07	.36**
		(.29)	(.74)	
Stability	42	2.31	2.54	−.22
		(.45)	(1.09)	

* $p < .05$.
** $p < .01$.

Research Question 4

In table 5, we report the results of our OLS regression that predicts standards-emphasized instruction based on the policy attribute gaps. We use the absolute value of the difference between paired teachers and principals to develop a gap variable to test our hypothesis that teachers and principals who experience policy similarly will have more standards-emphasized instruction in their schools. In other words, if teachers and principals are on the same page, standards-emphasized instruction is more likely to occur in a school, and teachers are more likely to feel prepared to teach all students. We hypothesized that gaps in the policy attributes between these two groups would be negative predictors for both instruction and preparedness.

For standards-emphasized instruction as an outcome, we have three significant finding concerning teacher-principal differences. We find that in Texas, for elementary ELA teachers, an authority gap is significant and negative. The R^2 value for Texas is also higher than the other states, suggesting that policy in Texas may be more tightly linked to instruction (the model explains standards-emphasized instruction better in Texas than in Ohio or Kentucky). Thus, Texas elementary ELA teachers who differ strongly from their principals in terms of belief in the standards teach significantly less of the emphasized content.

We find two significant positive predictors, again only in Texas. Gaps in specificity and power between teachers and principals are significant positive predictors of standards-emphasized instruction in ELA (either elementary or secondary). Remember that in Texas teachers and principals, on average, do not experience significant differences in specificity. But teachers do experience

TABLE 5

Regression Analysis of Standards-Emphasized Instruction with Attribute Gaps

	TEXAS				OHIO				KENTUCKY			
	ELA		Math		ELA		Math		ELA		Math	
GAP	Elem	High	Elem	High	Elem	High	Elem	High	Elem	High	Elem	High
Specificity	.27**	.06	−.06	−.11	−.04	−.10	−.19	−.22	−.20	−.30	−.39	.11
	(.08)	(.12)	(.13)	(.17)	(.12)	(.10)	(.20)	(.18)	(.22)	(.21)	(.30)	(.16)
Consistency	.16	.18	−.04	.31	−.10	−.07	.04	.38	−.02	.18	−.06	−.05
	(.13)	(.14)	(.19)	(.29)	(.19)	(.16)	(.32)	(.26)	(.20)	(.19)	(.33)	(.16)
Authority	−.25**	−.18	.04	.14	−.14	−.02	.18	.14	.06	.07	.21	−.05
	(.08)	(.17)	(.13)	(.24)	(.17)	(.11)	(.28)	(.17)	(.15)	(.13)	(.26)	(.12)
Power	−.08	.20*	.10	.07	.06	.11	−.23	−.17	.08	.09	−.01	.22
	(.07)	(.09)	(.11)	(.15)	(.12)	(.09)	(.18)	(.15)	(.14)	(.16)	(.27)	(.15)
All	−.01	−.07	.03	−.07	.03	.01	.06	−.03	.03	.03	.09	.01
	(.03)	(.04)	(.05)	(.07)	(.04)	(.04)	(.08)	(.06)	(.10)	(.10)	(.15)	(.07)
Constant	3.40**	3.55**	3.17**	3.18**	3.38**	3.48**	3.31**	3.10**	3.21**	2.93**	3.19**	2.85**
	(.09)	(.14)	(.15)	(.21)	(.17)	(.11)	(.28)	(.18)	(.28)	(.29)	(.46)	(.24)
n	70	53	73	34	44	50	47	40	35	32	24	42
R^2	.29	.22	.02	.07	.05	.05	.09	.17	.04	.13	.16	.14

NOTE.—Standard errors in parentheses. ELA = English language arts; Elem = elementary school; High = high school.

* $p < .05$.

** $p < .01$.

significantly more power than principals. Thus, if this power gap is predictive of standards-emphasized instruction, rewards and sanctions may be able to shift the instructional needle, even if these accountability measures do fall more heavily on teachers than principals.

When examining teacher preparedness as an outcome in table 6, we find one significant result in one state. In Ohio, teachers who do not agree with their principals on the alignment of curriculum, assessments, and professional development (consistency) are significantly less likely to feel prepared to teach the standards to all students.

Discussion

This survey analysis shows differing implementation of CCR standards and moderate buy-in of CCR standards across three states. One finding is unequivocal and surprisingly consistent across all three states: teachers do experience significantly more rewards and sanctions than principals and district administrators, and they show significantly weaker authority (buy-in) toward the standards. These findings mirror Desimone's (2006) study of an earlier wave of standards-based reform. Here is evidence that common perceptions of accountability policy persist. This finding also reflects a contemporaneous national poll of teachers, which finds teachers to be much less supportive of standards and testing than the general public (Henderson et al. 2017). Our study's unique design enables more targeted, state-level conclusions and, unlike Desimone's (2006) earlier cross-state standards implementation analysis, we find different patterns of relationships across states. Though relationships between policy and instruction may not be strong in every state (or surprising, considering the history of standards implementation), there is a continuing disconnect between teachers and administrators that holds serious policy implications. Along with Desimone (2006), our work is unique in its comparison of attitudes among three groups; few other studies leverage stratified surveys to quantify long-held assumptions about the implementation of standards-based reform or attempt to relate these policy attitudes to instruction.

We expect that groups that advocate for greater teacher autonomy, such as teachers unions, will remain energized on the issue of test-based accountability based on these power and authority gaps. A pressing question worth further study is why Texas policy appears to have stronger relationships with instruction than Ohio and Kentucky—two states that, unlike Texas, adopted the CCSS and participated in federalized initiatives such as Race to the Top (McGuinn 2012). Our qualitative work suggests that Texas's resistance to federal intervention may have offered more buy-in for state-level policies compared with states that adopted the CCSS (Desimone et al. 2019).

TABLE 6

Regression Analysis of Teacher Preparedness with Attribute Gaps

Gap	Texas	Ohio	Kentucky
Specificity	.0253	−.00715	.0926
	(.29)	(−.07)	(.79)
Consistency	−.0789	−.344*	.0746
	(−.62)	(−2.18)	(.64)
Authority	−.153	.0419	.161
	(−1.64)	(.32)	(1.85)
Power	.0145	.0766	.0716
	(.20)	(.82)	(.79)
Overall	.00231	−.0361	−.0104
	(.07)	(−.92)	(−.21)
Constant	3.472**	3.061**	2.842**
	(33.03)	(24.05)	(16.99)
n	141	106	97
R^2	.026	.088	.057

NOTE.—Standard errors in parentheses.
* $p < .05$.
** $p < .01$.

Some gaps are to be expected in the implementation of any reform, but this article confirms a popular assumption with empirical evidence. Teachers consistently report experiencing evaluative policies (reward and sanctions) more than principals and district officials, irrespective of different state policy contexts. Only recently have states begun to turn their attention to their principal evaluation systems (Donaldson and Woulfin 2018). Measures such as merit pay and value-added evaluations typically focus on teachers; a primary goal of standards-based reform is to align teacher instruction and raise student achievement (Smith and O'Day 1991).

But to achieve this goal, should teachers feel more accountability pressures than principals do, or should principals feel similar evaluative pressures? In Texas, a power gap between teachers and principals significantly predicts more standards-emphasized instruction, but it has no relationship to whether teachers feel prepared to help all students achieve grade-level standards in table 6. In Ohio, a consistency gap (alignment of policies and materials) between teachers and principals significantly predicts teachers feeling less prepared. Having aligned materials, curriculum, and professional development relates not only to teacher preparedness but also to standards-emphasized instruction. These relationships suggest that policy makers and researchers should pay close attention to teachers' perceptions of policy if standards-based reform is to be fully realized and well implemented.

Implementation itself is like "the human telephone game" (Spillane 2009, 8), with the best-intentioned often misinterpreting a policy's intent. If teachers feel they are being punished or rewarded more heavily than principals, different incentive structures may exist within the same building and at cross-purposes with each other. Our authority and power findings speak to an ongoing challenge between teachers on the ground and those evaluating the implementation of standards from central offices or from the main office just down the hallway. These findings suggest a policy environment that is loosely coupled (Weick 1976) except in Texas, where authority significantly predicts standards-emphasized instruction (Edgerton and Desimone 2018).

The lack of significance around specificity may be a testament to progress made in sharing resources that are aligned with similar state standards. But when we dig deeper and compare teachers with district officials, we find that central offices believe that they are providing significantly more specificity than teachers report receiving. Considering state averages, teachers are not as concerned about alignment with the standards (consistency) as they were in prior periods of reform (Desimone 2002, 2006). We find relatively high consistency scores and few significant relationships. Shared curricular resources across state lines and widely respected programs such as Eureka Math may have contributed to this improvement in consistency, though of course this is only one possible explanation (Hodge et al. 2016).

Limitations and Considerations in Interpreting the Results

A major limitation of this study is that these are findings correlational, and we did not collect extensive demographic information about respondents that would control for respondent characteristics (with the exception of years of teaching experience). We acknowledge well-documented issues with survey self-report and followed guidance on developing valid behavior-based questions rather than evaluative questions (Desimone 2006; Mayer 1999). For example, we did not signify which content items were aligned with the standards. Thus, teachers were not evaluating specific practices but rather were asked to assign relative emphasis to content on a sliding scale. Furthermore, teachers reported on a range of subject-specific content, and EL and SWD teachers reported on both subjects, which should make our instructional conclusions more valid. Our method (Surveys of Enacted Curriculum) has also been shown to be reliable and valid across multiple studies (Blank 2002; Osthoff 2007; Polikoff et al. 2011; Rowan et al. 2004; Webb 2002, 2007).

Another important consideration is that the quality of teaching can differ dramatically from observation and self-report (Cohen 1990), and quality and

instructional content may not be closely related (Polikoff and Porter 2014). Triangulating our data with additional policy document analysis and classroom observations would undoubtedly add dimensions and insight to our findings.

Conclusion

With accountability systems now changing under approved state ESSA plans, important questions remain about the appropriate role of rewards and sanctions targeted toward teachers. If accountability can be successfully tied to improved outcomes as Dee and Jacob (2011) and others have done in earlier policy periods, perhaps these policies may find renewed political and public support. But as of spring 2016 in Texas, Ohio, and Kentucky, those ostensibly in charge of whole-school reform—namely, principals—feel significantly weaker rewards and sanctions than teachers in their classrooms. Simultaneously, most state and district officials believe that they are weakening their systems of rewards and sanctions (Desimone et al. 2019). In addition, it is hard to argue that significantly lower authority (buy-in) for teachers around CCR standards is a positive development. This state of affairs suggests a ground for future tension and negotiation around who bears what share of accountability and which actors are driving the next iteration of standards-based reform.

We can neither prescribe definitive explanations based on our related qualitative work nor tie survey results to specific state policies. But we can contest a common theme found in our interviews with state and district officials—namely, that they have disavowed NCLB-style accountability (Desimone et al. 2019). In fact, in every state, teachers report experiencing significantly more power than principals or superintendents regardless of policy differences across states. Recurring patterns across states may demonstrate that teachers are focused on their students, not on specific state policies.

Still, we can we point to Texas as an important state for future study. It may have policy structures worth emulating. Or it may have invested more energy up front in developing the TEKS and in implementing them more slowly and carefully. We can speculate that our standards-emphasized instruction model explains more of the variance in Texas because CCR standards there did not begin with the CCSS. The rejection of the CCSS may have been a pivotal decision, and we urge researchers to look more closely at other CCSS rejecters. It seems, however, that there is a universality to teacher experience with standards-based policy, one that may be entrenched by two decades of NCLB-era testing. Despite movement away from power mechanisms and toward building authority (Desimone et al. 2019), teachers feel the repercussions of accountability policy significantly more than principals or district officials do. This disproportionality, for all of its political costs, positively relates to instruction in only one

subject in one grade and one state. Principals and superintendents, not teachers, may be the better targets for increased accountability.

Notes

This research was supported in part by Grant R305C150007 from the Institute of Education Sciences in the U.S. Department of Education to the University of Pennsylvania.

1. "District officials" refers to superintendents or those individuals with similar positional authority within a school district's central office who answered the survey.
2. Throughout this article, we use the term "CCR standards" to refer to each state's current standards in math and English language arts (ELA). These are best understood as being similar to the Common Core State Standards after individual state modification.
3. Copies of the full teacher surveys are provided at https://www.c-sail.org/resources.

References

Berends, Mark, Susan J. Bodilly, and Sheila N. Kirby. 2002. *Facing the Challenges of Whole-School Reform: New American Schools after a Decade*. Santa Monica, CA: RAND, https://www.rand.org/pubs/monograph_reports/MR1498.html.

Blank, Rolf K. 2002. "Using Surveys of Enacted Curriculum to Advance Evaluation of Instruction in Relation to Standards." *Peabody Journal of Education* 77 (4): 86–121.

Boyd, Donald, Pam Grossman, Marsha Ing, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2011. "The Influence of School Administrators on Teacher Retention Decisions." *American Educational Research Journal* 48 (2): 303–33.

Burkhauser, Susan. 2017. "How Much Do School Principals Matter When It Comes to Teacher Working Conditions?" *Educational Evaluation and Policy Analysis* 39 (1): 126–45.

Carmichael, Sheila B., Gabrielle Martino, Kathleen Porter-Magee, and W. Stephen Wilson. 2010. *The State of State Standards—and the Common Core—in 2010*. Washington, DC: Fordham Institute.

Clune, William H. 1993. "Systemic Educational Policy: A Conceptual Framework." In *Designing Coherent Educational Policy*, ed. S. H. Fuhrman. San Francisco: Jossey-Bass.

Coburn, Cynthia E., and Joan E. Talbert. 2006. "Conceptions of Evidence Use in School Districts: Mapping the Terrain." *American Journal of Education* 112 (4): 469–95.

Cohen, David K. 1990. "A Revolution in One Classroom: The Case of Mrs. Oublier." *Educational Evaluation and Policy Analysis* 12 (3): 327–45.

Cohen, David K., and Jal Mehta. 2017. "Why Reform Sometimes Succeeds: Understanding the Conditions That Produce the Reforms That Last." *American Educational Research Journal* 54 (5): 644–90.

C-SAIL (Center on Standards, Alignment, Instruction, and Learning). 2018. "State College- and Career-Readiness Policies and the Features That Make Their Implementation Successful." https://www.c-sail.org/maps/standards-policies-features.

C-SAIL (Center on Standards, Alignment, Instruction, and Learning). 2019. "State Accountability Systems and the Policies That Support Them." https://www.c-sail.org/maps/accountability.

Dee, Thomas S., and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and Management* 30 (3): 418–46.

Desimone, Laura M. 2002. "How Can Comprehensive Reform Models Be Successfully Implemented?" *Review of Educational Research* 72 (3): 433–79.
Desimone, Laura M. 2006. "Consider the Source: Response Differences among Teachers, Principals and Districts on Survey Questions about Their Education Policy Environment." *Educational Policy* 20 (4): 640–76.
Desimone, Laura M. 2009. "Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures." *Educational Researcher* 38 (3): 181–99.
Desimone, Laura M., and Kerstin Carlson Le Floch. 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.
Desimone, Laura M., Amy Stornaiuolo, Nelson Flores, Adam K. Edgerton, T. Phillip Nichols, Katie Pak, Emily Plummer, and Andrew C. Porter. 2019. "Success and Challenges of the 'New' College- and Career-Readiness Standards: Seven Implementation Trends." *Educational Researcher* 48 (3): 167–78.
Desimone, Laura M., Tonya Wolford, and Kirsten Lee Hill. 2016. "Research-Practice: A Practical Conceptual Framework." *AERA Open* 2 (4): 1–14.
DiMaggio, Paul, and Walter W. Powell. 1983. "The Iron Cage Revisited: Collective Rationality and Institutional Isomorphism in Organizational Fields." *American Sociological Review* 48 (2): 147–60.
Donaldson, Morgaen L., and Sarah Woulfin. 2018. "From Tinkering to Going 'Rogue': How Principals Use Agency When Enacting New Teacher Evaluation Systems." *Educational Evaluation and Policy Analysis* 40 (4): 531–56.
Dunn, Alyssa H., Matthew Deroo, and Jennifer VanDerHeide. 2017. "With Regret: The Genre of Teachers' Public Resignation Letters." *Linguistics and Education* 38 (1): 33–43.
Edgerton, Adam K., and Laura M. Desimone. 2018. "Teacher Implementation of College- and Career-Readiness Standards: Links among Policy, Instruction, Challenges, and Resources." *AERA Open* 4 (5): 1–22.
Edgerton, Adam K., Douglas Fuchs, and Lynn Fuchs. Forthcoming. "New Standards and Old Divides: Policy Attitudes about College- and Career-Readiness Standards for Students with Disabilities." *Teachers College Record*.
Flores, Nelson, and Jamie L. Schissel. 2014. "Dynamic Bilingualism as the Norm: Envisioning a Heteroglossic Approach to Standards-Based Reform." *TESOL Quarterly* 48 (3): 454–79.
Foucault, Michael, and Alan Sheridan. 1975. *Discipline and Punish: The Birth of the Prison*. New York: Random House.
Gamson, David A. 2015. "Differences in Common." In *Challenging Standards: Navigating Conflict and Building Capacity in the Era of the Common Core*, ed. J. P. Spillane and J. A. Supovitz. Lanham, MD: Rowman & Littlefield.
Goldring, Ellen, and Patrick Schuermann. 2009. "The Changing Context of K–12 Education Administration: Consequences for EdD Program Design and Delivery." *Peabody Journal of Education* 84 (1): 9–43.
Griffith, James. 2004. "Relation of Principal Transformational Leadership to School Staff Job Satisfaction, Staff Turnover, and School Performance." *Journal of Educational Administration* 42 (3): 333–56.
Gutiérrez, Rochelle, and Ezekiel Dixon-Román. 2011. "Beyond Gap Gazing: How Can Thinking about Education Comprehensively Help Us (Re)envision Mathematics Education?" In *Mapping Equity and Quality in Mathematics Education*, ed. B. Atweh, M. Graven, W. Secada, and P. Valero. Dordrecht: Springer.

Hargreaves, Andy. 2010. "Presentism, Individualism, and Conservatism: The Legacy of Dan Lortie's *Schoolteacher*: *A Sociological Study*." *Curriculum Inquiry* 40 (1): 143–54.
Henderson, Michael B., Martin R. West, Paul E. Peterson, and Samuel Barrows. 2017. "Ten- Year Trends in Public Opinion from the EdNext Poll." *Education Next* 17 (1): 8–28.
Hodge, Emily M. 2018. "'Common' Instruction? Logics of Ability and Teacher Decision Making across Tracks in the Era of Common Standards." *American Educational Research Journal*. https://doi.org/10.3102/0002831218803328.
Hodge, Emily M., Serena J. Salloum, and Susanna L. Benko. 2016. "(Un)commonly Connected: A Social Network Analysis of State Standards Resources for English Language Arts." *AERA Open* 2 (4): 1–19.
Honig, Meredith I. 2006. "Street-Level Bureaucracy Revisited: Frontline District Central Office Administrators as Boundary Spanners in Education Policy Implementation." *Educational Evaluation and Policy Analysis* 28 (4): 357–83.
Honig, Meredith I. 2008. "District Central Offices as Learning Organizations: How Sociocultural and Organizational Learning Theories Elaborate District Central Office Administrators' Participation in Teaching and Learning Improvement Efforts." *American Journal of Education* 114 (4): 627–64.
Honig, Meredith I. 2009. "No Small Thing: School District Central Office Bureaucracies and the Implementation of New Small Autonomous Schools Initiatives." *American Educational Research Journal* 46 (2): 387–422.
Ingersoll, Richard, Lisa Merrill, and Daniel Stuckey. 2014. *Seven Trends: The Transformation of the Teaching Force*. CPRE Research Report RR–80. Philadelphia: Consortium for Policy Research in Education.
Kaufman, Julia H., Jill S. Cannon, Shelly Culbertson, Margaret Hannan, Laura S. Hamilton, and Sophie Meyers. 2018. *Raising the Bar: Louisiana's Strategies for Improving Student Outcomes*. Santa Monica, CA: RAND.
Kaufman, Julia H., V. Darleen Opfer, Michelle Bongard, and Joseph D. Pane. 2018. *Changes in What Teachers Know and Do in the Common Core Era: American Teacher Panel Findings from 2015 to 2017*. Santa Monica, CA: RAND.
Kentucky Department of Education. 2019. "School-Based Decision Making." https://education.ky.gov/districts/SBDM/Pages/default.aspx.
Ladd, Helen F. 2017. "No Child Left Behind: A Deeply Flawed Federal Policy." *Journal of Policy Analysis and Management* 36 (2): 461–69.
Mayer, Daniel P. 1999. "Measuring Instructional Practice: Can Policymakers Trust Survey Data?" *Educational Evaluation and Policy Analysis* 21 (1): 29–45.
McGuinn, Patrick. 2012. "Stimulating Reform: Race to the Top, Competitive Grants and the Obama Education Agenda." *Educational Policy* 26 (1): 136–59.
Mehta, Jal. 2013. *The Allure of Order: High Hopes, Dashed Expectations, and the Troubled Quest to Remake American Schooling*. New York: Oxford University Press.
National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform: A Report to the Nation and the Secretary of Education*. Washington, DC: US Department of Education.
Ohio Department of Education. 2019. "Adoption and Purchase of Textbooks." http://education.ohio.gov/Topics/Learning-in-Ohio/Adoption-and-Purchase-of-Textbooks.
Osthoff, Eric. 2007. *Using SEC Data to Improve Curriculum Content Alignment in Oregon's Canyon Falls School District: Initial Steps and Evolving Challenges*. Madison, WI: Center for Education Research.
Pogodzinski, Ben, Peter Youngs, Kenneth A. Frank, and Dale Belman. 2012. "Administrative Climate and Novices' Intent to Remain Teaching." *Elementary School Journal* 113 (2): 252–75.

Polikoff, Morgan S. 2012. "The Association of State Policy Attributes with Teachers' Instructional Alignment." *Educational Evaluation and Policy Analysis* 34 (3): 278–94.
Polikoff, Morgan S., Tenice Hardaway, Julie A. Marsh, and David N. Plank. 2016. "Who Is Opposed to Common Core and Why?" *Educational Researcher* 45 (4): 263–66.
Polikoff, Morgan S., and Andrew C. Porter. 2014. "Instructional Alignment as a Measure of Teaching Quality." *Educational Evaluation and Policy Analysis* 36 (4): 399–416.
Polikoff, Morgan S., Andrew C. Porter, and John Smithson. 2011. "How Well Aligned Are State Assessments of Student Achievement with State Content Standards?" *American Educational Research Journal* 48 (4): 965–95.
Porter, Andrew C. 1994. "National Standards and School Improvement in the 1990s: Issues and Promise." *American Journal of Education* 102 (4): 421–49.
Porter, Andrew C. 2002. "Measuring the Content of Instruction: Uses in Research and Practice." *Educational Researcher* 31 (7): 3–14.
Porter, Andrew C., Robert Floden, Donald Freeman, William Schmidt, and John Schwille. 1988. "Content Determinants in Elementary School Mathematics." In *Perspectives on Research on Effective Mathematical Teaching*, ed. D. A. Grouws and T. J. Cooney. Hillsdale, NJ: Erlbaum.
Powell, Sarah R., Lynn S. Fuchs., and Doug Fuchs. 2013. "Reaching the Mountaintop: Addressing the Common Core Standards in Mathematics for Students with Mathematics Difficulties." *Learning Disabilities Research and Practice* 28 (1): 38–48.
Rowan, Brian, Eric Camburn, and Richard Correnti. 2004. "Using Teacher Logs to Measure the Enacted Curriculum: A Study of Literacy Teaching in Third-Grade Classrooms." *Elementary School Journal* 105 (1): 75–101.
Satz, Debra. 2007. "Equality, Adequacy, and Education for Citizenship." *Ethics* 117 (4): 623–48.
Smith, Marshall S., and Jennifer O'Day. 1991. *Putting the Pieces Together: Systemic School Reform*. New Brunswick, NJ: Consortium for Policy Research in Education.
Spillane, James P. 2009. *Standards Deviation: How Schools Misunderstand Education Policy*. Cambridge, MA: Harvard University Press.
Stosich, Elizabeth Leisy. 2016. "Joint Inquiry: Teachers' Collective Learning about the Common Core in High-Poverty Urban Schools." *American Educational Research Journal* 53 (6): 1698–731.
Supovitz, Jonathan. 2017. "Social Media Is the New Player in the Politics of Education." *Phi Delta Kappan* 99 (3): 50–55.
Supovitz, Jonathan, Philip Sirinides, and Henry May. 2009. "How Principals and Peers Influence Teaching and Learning." *Educational Administration Quarterly* 46 (1): 31–56.
Tampio, Nicholas. 2016. "Democracy and National Education Standards." *Journal of Politics* 79 (1): 33–44.
Texas Education Agency. 2019. "Texas Essential Knowledge and Skills." https://tea.texas.gov/curriculum/teks.
Webb, Norman. 2002. *Alignment Study of Language Arts, Mathematics, Science, and Social Studies of State Standards and Assessments in Four States: A Study of the State Collaborative on Assessment and Student Standards (SCASS), Technical Issues in Large- Scale Assessment (TILSA)*. Washington, DC: Council of Chief State School Officers.
Webb, Norman. 2007. "Issues Related to Judging the Alignment of Curriculum Standards and Assessments." *Applied Measurement in Education* 20 (1): 7–25.
Weick, Karl E. 1976. "Educational Organizations as Loosely Coupled Systems." *Administrative Science Quarterly* 21 (1): 1–19.