Successes and Challenges of the "New" College and Career Ready Standards: Seven Implementation Trends

Published April 2019 in *Educational Researcher, 48*(3), 167-178.

Laura M. Desimone, Amy Stornaiuolo, Nelson Flores, Katie Pak, Adam Edgerton, T. Philip Nichols, Emily C. Plummer, Andrew Porter

Abstract

This study identifies seven major trends in how states and districts are implementing college- and career-ready standards for general education students and for two special populations often the target of education policy—English language learners (ELLs) and students with disabilities (SWDs). We draw on state-representative teacher, principal, and district surveys in three states—Kentucky, Ohio, and Texas—and case studies in nine districts. We ground our study in the policy attributes framework, which suggests implementation is stronger the more specific, authoritative, powerful, consistent, and stable a policy is. We find states are being less prescriptive in their policies surrounding the standards and are including fewer or less forceful rewards and sanctions (power). Local districts are providing more detailed, standards-aligned professional development (specificity) and supporting materials to guide teachers' standards implementation (consistency). Districts are using "softer" power mechanisms instead of the "strong" rewards and sanctions of earlier waves of reform. This results in higher buy-in (authority) but creates challenges for districts in providing the necessary supports for teachers. In ELL policy, two national organizations are providing much of the specificity and consistency for standards implementation, and they do this through mechanisms of authority rather than through power mechanisms. For SWDs, implementation support is focused on compliance, and the enduring tension between standardization and individuality persists. Creative district approaches and moderate to high levels of authority hold promise for this wave of college- and career-ready standards.

Keywords: accountability; educational policy; educational reform; mixed methods

For nearly four decades, educators and policy makers have sought to improve teaching and learning by establishing content standards that serve as the foundation for K–12 curricula and instruction and by developing aligned assessments. A new wave of standards-based reform followed the 2010 adoption of the Common Core State Standards (CCSS), with the goal of ensuring that students are prepared to succeed in college and careers. While the CCSS were initially adopted by virtually all states, many have since replaced them with their own standards for college and career readiness. As called for by the Every Student Succeeds Act of 2015, all 50 states have now adopted new and more challenging content standards for K–12 education. Here we examine this latest wave of reform efforts.

This analysis is part of a larger research project undertaken by the Center on Standards, Alignment, Instruction, and Learning (C-SAIL), an IES funded center investigating the implementation and effects of college- and career-ready standards for all students, including English language learners (ELLs) and students with disabilities (SWDs).

Our research is grounded in the policy attributes framework, which posits that the more specific, consistent, authoritative, powerful, and stable a policy is, the better its implementation will be (Porter, 1994; Porter, Floden, Freeman, Schmidt, & Schwille, 1988).

We draw on data from three states—state-representative surveys of 84 district officials, 439 principals and 1,760 teachers, 36 interviews with state officials and 54 district interviews across nine case study districts. We chose our three study states—Kentucky, Ohio, and Texas—to vary on whether they had their own standards or Common Core standards, implementation timelines, rigor of standards, urbanicity, and ethnic composition, as well as their willingness to partner with us. We studied each state's standards—the Kentucky Academic Standards (KAS), Ohio's Learning Standards, the Texas Essential Knowledge and Skills (TEKS), and the state and district descriptions of the policies around how they are implemented.

The three states have widely divergent standards-based education policies with different implementation timelines. Texas, a state dominated by a single party, has maintained a remarkably stable and consistent policy environment, eschewing the CCSS and charting its own path. The other two states, by contrast, have been subject to much more political instability, with parties switching control of the legislature and governorship, resulting in both states adopting and then repealing the CCSS. In terms of alignment, Kentucky mitigated some of this instability by relying on its own state-developed assessment, just as Texas did, whereas Ohio designed a new assessment after withdrawing from PARCC and using interim assessments. Of the three states, Ohio has the strongest teacher accountability policies. In terms of local control over curricula, Kentucky ranks highest with its use of School-Based Decision Making Councils, while both Texas and Ohio use state committees to approve aligned curricula and textbooks. Texas's standards were significantly more rigorous than the Ohio and Kentucky standards prior to adoption of the CCSS (Carmichael, Martino, Porter-Maggee, & Wilson, 2010). 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 and Ohio's K-12 students are still overwhelmingly White, which is not the case in Texas. Considering these important differences (which would make similarities all the more remarkable), we ask the following questions:

- What are major trends in how these states and districts are implementing college- and career-ready standards?
- What are the major implementation trends for teachers of English language learners and students with disabilities?

We found that states are leaving much more of the standards implementation work to local districts compared with standards-based reform efforts from the mid-1990s to early 2010s. States are being less prescriptive and detailed in their policies surrounding the standards (the policy attribute of

specificity), and they are including fewer or less forceful rewards and sanctions (the policy attribute of power). Local districts are stepping into the policy space that states have vacated by developing more detailed, standards-aligned professional development and supporting materials to guide teachers in implementing the standards (the policy attributes of specificity and consistency). Our data also indicate that districts are using "softer" power mechanisms instead of emphasizing the kinds of "strong" rewards and sanctions that were part of earlier waves of reform (Cohen & Mehta, 2017; Mehta, 2013).

These local efforts are resulting in higher teacher buy-in for the standards, which bodes well for the success of this round of reform. One potential challenge, though, is that many districts—especially small, under-resourced, rural districts—are struggling to undertake the efforts once handled by the state.

In contrast, for standards implementation for ELLs and SWDs, we found that the state plays a more visible role. The states, in partnership with national organizations, seem to be providing much more specificity and consistency (aligning supportive materials and PD with standards and assessments) in implementation standards for ELLs. They do this through authority mechanisms that give the standards legitimacy and generate buy-in (e.g., rules, historical practice) rather than through power mechanisms (e.g., rewards and sanctions). In terms of standards implementation for SWDs, however, the state plays a central role in ensuring compliance with federal policy through specific, consistent, and powerful standards implementation rather than by building authority. This trend contributes to the enduring tension between standardization and individuality in the implementation of standards for SWDs, which continues to create challenges for educators.

In Theory There Is No Difference Between Theory and Practice. In Practice There Is[1]**:**

[1] Section titles in italics thanks to Yogi Berra.

The Theory Guiding Our Study

As noted above, the policy attributes theory posits that the more specific, consistent, authoritative, powerful, and stable a policy is, the better the implementation will be. *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, buy-in, 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., curricula, assessments, professional development, and evaluations). *Power* reflects how policies are reinforced and enacted through systems of rewards (e.g., incentives for compliance) and sanctions (e.g., dismissals for low performance on tests). *Stability* refers to the extent to which policies change or remain constant over time (e.g., how long standards and assessments have been in place).

The policy attributes theory has been used for decades in education policy research. It was developed for a study analyzing how teachers made determinations about the content they taught in class (Porter et al., 1988), and as such provided a relevant framework as teachers' content decisions became increasingly influenced by the national standards movement (Porter, 1994). Subsequently it has been applied to understand the components of systemic reform (Clune, 1993), and to describe the implementation and effects of comprehensive school reform (Berends, Bodilly, & Kirby, 2002; Desimone, 2002). Recently, the theory has been used to categorize state policy and its relationship to aligned instruction (Polikoff, 2012) and to analyze the implementation and effects of a research-practice partnership (Desimone, Wolford & Hill, 2016).

Methods: *You Can Observe a Lot Just by Watching* (and Listening and Asking)

The data for this article are from surveys and interviews conducted in the spring and summer of 2016 and 2017 to allow for both triangulation and complementarity (Ravitch & Carl, 2016). Table 1 describes the number of teachers, principals, district and state officials in our survey

and interview samples. The survey study has a multistage sampling design, with districts selected with probability proportional to the square root of student enrollment size, with schools selected with equal probability within the elementary school group and the high school group, and with teachers selected with equal probability within each teacher group. We employed a stratified random sampling technique designed to ensure the sample was representative of districts in Texas, Ohio, and Kentucky. We included 42 Texas districts, 42 Ohio districts, and 89 Kentucky districts in the sample (there was a larger number in Kentucky because of integration with state-sponsored survey data collection which included the universe of districts). 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. 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 ELLs. In each high school, we sampled two English Language Arts teachers and one teacher in each of the following specialties or topics: SWDs, ELLs, algebra I, algebra II, and geometry. To correct for any response bias, we used survey jackknife procedures in Stata and report robust standard errors throughout.

[Insert Table 1 about here]

Our interview study was embedded in the state-representative survey study; we conducted structured interviews, lasting 30 to 90 minutes each, with state officials and district officials in three districts in each state—one urban, one suburban, and one rural, to allow the exploration of geographic patterns in our data. We targeted interview districts to have a population of SWDs and ELLs representative of the average in the state (e.g., Texas districts on average have much higher numbers of ELLs than Ohio districts). Table 2 indicates the SWD and ELL populations in each of our case study districts, by urbanicity. We chose respondents who were in the support and implementation of college- and career-ready standards in their state or district.

[Insert Table 2 about here]

The purpose of our survey and interviews was to learn about the quality of implementation of standards, and barriers and facilitators to implementation, grounded in the policy attributes conceptual framework. We had a core of questions that were the same across role groups, to allow comparison of different perspectives, in addition to a set of questions that applied only to specific roles. Specifically, we asked districts questions about state policy; we asked principals questions about district resources, supports and policy; and we asked teachers about school resources, supports and policy, as well as classroom-level challenges and affordances. In the Appendix we provide our interview questions and a selection of our survey questions with a link to the full survey. Survey responses ranged from 1-4, where 1 = strongly disagree, 2 = somewhat disagree, 3 = somewhat agree, and 4 = strongly agree.

The survey analyses we report are based on descriptive analyses. When comparing different groups of teachers, and when comparing teachers to principals to district administrators, we used jackknife resampling, which provides robust standard errors and allows for Wald tests of significance across the groups. Throughout the paper, we use the averages obtained from jackknife resampling to correct for any nonresponse bias. When comparing states to each other, we used an ANOVA test of the summary results obtained from the jackknife procedure with a Tukey post-hoc correction to observe which of the three states were statistically significant from each other.

Transcribed interviews were analyzed using multiple rounds of inductive and deductive coding. In the first round of analysis, we developed coding categories based on the policy attributes as well as key reform areas (e.g., curriculum, instruction, PD), followed by second-round analyses that added emergent codes (e.g., governance, communication strategies). During analysis of the coded transcripts, we created a data matrix to determine how the policy implementation categories (e.g., professional development, curriculum) interacted with the policy attributes (e.g., authority,

specificity). We identified categories that appeared across multiple sectors of the data matrix, such as the theme of 'local control' that appeared in every category for every state. We generated cross-cutting themes from these categories, going back to the transcripts to identify evidence from the transcripts for these themes, while simultaneously refining the themes based on this evidence, and looking for any evidence contrary to the emerging themes. We developed inter-rater reliability in coding through a process of paired coding, research team discussion, and recoding. Our interview protocols and coding system are available in the Appendix.

In reporting our findings, we follow Atkinson, Coffey & Delamont (2003), Coffey & Atkinson (1996), and Ryan & Bernard (2003) in the determination of themes, and using illustrative quotes to represent those themes. Our analytic process included inquiry into dissensus in the data; however, one of the findings most compelling to us was that the seven trends we identify were present in some way within district and state levels, and across our three study states, with no evidence of disagreement in our data.

Trend #1

***Déjà vu All Over Again*: The Pendulum Is Swinging Back To Local Control—at Least for General Education**

The balance of power between states and localities has been an issue in education since the early 1900s (Tyack, 1974), with approaches toward standards implementation varying in the degrees to which states controlled the overarching goal and vision and localities determined the means (Spillane, 2009). While norms around local control have remained strong through the decades (Grissom & Herrington, 2012), during No Child Left Behind (NCLB), states exercised considerable control over standards and accountability policy (Wong et al., 2017). However, in this most recent wave of standards reform, we find the pendulum swinging back to local control, with local actors playing a major role in providing specific support and guidance on implementing standards. This is

consistent with research showing stronger local roles in other realms of education policy (Marsh & Wohlstetter, 2013).

Approximately 75% of principals in all three of our study states reported on a survey that they provide guidance to teachers on how much time to spend on content and the order in which it should be taught. And across all three states, teachers and principals agreed that their district provided specific guidelines for implementing their state's standards (Edgerton & Desimone, 2018).

District interview respondents in Kentucky, Ohio, and Texas all indicated that their state department of education provides support and guidance for standards implementation in the form of model curricula, alignment maps, and other similar resources, but that ultimately districts and schools play a major role in identifying and providing the specific supports for implementation. The majority of districts indicated that they appreciated this autonomy of standards implementation, with a suburban district in Kentucky acknowledging that the process of providing their own materials "wasn't bad because it required us to dig into the standards" as a form of professional learning. As one district official in Texas similarly noted: "Of course the state pushes about college and career readiness, but I think it's more to the local entities to really standardize that, articulate it, and create thoughtful plans toward that. It's like the state provides you the overall goal, overall umbrella, but then it's up to the local districts to really implement more defined plans." This was considered a strength, reflecting research that shows the importance of being able to anticipate and calibrate reforms to the contextual complexities of particular schools and classrooms (Desimone & Hill, 2017).

Both state and district officials in all three states reported a key tension around respecting the need for district, school, and teacher autonomy in the implementation process while also providing enough specific and useful guidance to support standards implementation. Our survey results indicate that district leaders and principals, as compared with teachers, wanted significantly more information about how the standards changed what was expected of teachers (see Figure 1). Consistent with previous research that showed differences in how teachers, district officials and principals view the standards environment (Desimone, 2013) the differences between teachers and principals, and teachers and districts, was statistically significant at the .05 level. This reflects the idea that districts and school leaders view their job as helping teachers interpret the standards, while teachers believe they already understand the standards and mainly want more supports for implementing them (Edgerton & Desimone, 2018). High on the list of supports for all three groups was digital tools, an increasingly popular instruction support (Anglum & Desimone, forthcoming).

[Insert Figure 1 about here]

Trend #2

***When You Arrive at a Fork in the Road, Take It*: In Grappling with Alternative Approaches, Districts Struggle with Developing Specific, Aligned Resources and Establishing Specific and Consistent Implementation Practices Across Schools**

As control over standards implementation has shifted to districts, some are struggling to provide specific and aligned resources and supports for teachers. Whereas in previous waves of reform many districts purchased standards-aligned textbooks and curricula to ensure their classroom practices were reliably linked to state expectations and assessments (Gewertz, 2015), in our case studies, district leaders explained that there were a number of reasons districts had to find and develop these resources on their own: because resources were difficult to find on the state website, or the state provided only a few model units, or states avoided endorsing specific materials for fear

of being too prescriptive (or facing political backlash). In Texas, for example, an urban district official reported that the state provided a list of materials and indicated which materials were aligned to the state standards, but if they want a particular resource "[it's] up to you to go look for it."

The time and labor involved in creating aligned materials, particularly for districts that have comparatively fewer resources, represented a considerable investment. KY district officials described a desire for more aligned resources, with one suburban official reporting that "it would have been nice to have more support [from the state]," particularly around "more support within actual instruction. It seems like it's always about rules and procedures and how to do certain programs correctly versus supporting the actual instruction of the student." In both Ohio and Texas, the rural districts reported on the major challenges of developing aligned curriculum in small and under-resourced districts. They, like the Texas suburban district, felt like they were often left to "fend for themselves."

Further, in the wake of loosened state control, districts find it challenging to identify and address variation across schools in their approach to standards implementation. We found that increased school-based autonomy, which often accompanied local control culture, sometimes resulted in unproductive variation in implementation (e.g., redundant activities), which districts struggled to address. Commonly, professional development (PD) represents a critical mechanism for operationalizing accountability systems by communicating the standards and supporting consistent use (Hochberg & Desimone, 2010); however, due to union rules, some districts could not mandate particular PD activities (see Grissom & Strunk, 2010) and so were limited in using district-sponsored PD to help establish consistent practices across schools.

One likely effect of this lack of central control over PD is the considerable variation in the range of time teachers and principals spend in standards PD in math and literacy, ranging from 0 to more than 80 hours (see Table 3), with a skewness toward zero. More than two thirds of teachers

(68%) spent 20 hours or less on standards PD, which prior research suggests is not enough to foster meaningful change (Desimone, 2002; Desimone & Garet, 2015).

[Insert Table 3 about here]

Trend #3

***How Can You Think and Hit at the Same Time?* Schools Achieving Consistency & Specificity in Their Implementation Practices Offer More Professional Development**

While our descriptive statistics show low overall levels for specificity and hours for PD, we find that as consistency and specificity increase, so do PD hours, suggesting that districts may be using PD as a mechanism to achieve these policy attributes. Reform studies in the early 2000s reflect this idea that PD builds authority, and occurs more often in environments that develop consistency across reforms, and provide specific guidelines for implementing the reform (Desimone, 2002); our data show this seems to be the case with CCRS as well. In our case studies, all of our districts reported employing various forms of PD to provide guidance on implementing the standards (specificity) in ways that are well-aligned to the standards (consistency). These efforts include building school leader knowledge and skills to guide their teachers, instructional coaches to provide on-site assistance, and professional learning community (PLC) protocols to encourage systematic processes for reviewing teachers' work. District respondents described using instructional coaches for a range of functions related to standards implementation, such as providing content-specific PD, offering SWD and ELL support, and co-teaching and modeling instructional shifts. Coaching, which reflects many features of high-quality PD but has a mixed research base (Desimone & Pak, 2017) represents a considerable investment by districts, with 46% of teachers across all three states reporting participation in some form of coaching related to standards implementation.

The districts' investments in professional learning appear to add legitimacy (i.e., authority) to standards-based reform. One rural Texas official reflected on the substantial investment in coaching

the district has made and the resulting impact on teachers: "every single campus in our district has a campus instructional coach, and that person provides direct teacher coaching and support all year long and they're housed on the campus." Similarly, in Kentucky, district officials specifically referenced the role out of instructional rounds (where groups of teachers observe instruction and discuss what they observe) as their newest professional learning initiative that has been well-received: when district officials first introduced instructional rounds as a method for collaboratively understanding and implementing the standards, "you could see the principals' faces of 'oh my gosh this is what we've needed for a while." This investment in coaching is a way of "recogniz[ing] the incredible importance of professional development" that signals the authority of the new standards. In Ohio, district officials discussed using the outcomes of PLC meetings as the basis of their resource allocation decisions; this effort to use teacher input to guide district decision-making is another recognized way of building authority (Desimone, 2002; Desimone, Wolford, & Hill, 2016).

Trend #4

***The Future* (of Power) *Ain't What it Used to Be:* A Decrease in Punitive Sanctions at the State Level Has Fostered the Use of "Softer" Rewards & Incentives at the District Level**

We identified major shifts in how rewards and sanctions are communicated and deployed within and across states and districts. Assessments remain the primary means for determining accountability, but the weight of these assessments does not carry the same threat of punitive force that was well documented in previous waves of standards-based reform (Booher-Jennings, 2005; Darling-Hammond, 2004). While researchers and policymakers may see standards and accountability policy as separate, our respondents thought of accountability policy as a part of standards policy – specifically, that rewards and sanctions are part of the system that motivates educators to follow the standards.

Teachers' in all three states indicated that the rewards and sanctions associated with standards implementation—power in our framework—were moderate, in the 2.30-2.68 range, on a scale of 1 to 4 (see Table 4). We consider scores below 2 as low, since 1=strongly disagree and 2=disagree. Moderate scores fall between 2 and 3, where 3=agree and 4=strongly agree. Teachers reported experiencing significantly more power (rewards and sanctions) compared to principals in Texas and Kentucky, and compared to district administrators in Ohio. Even so, all respondents perceived low to moderate effects of power in standards implementation.

[Insert Table 4 about here]

Our interview data reflect a similar pattern, showing that softer power has eclipsed stronger forms. Without exception, state officials in all three states stressed that decisions about how to use assessment data in rewarding or sanctioning schools or teachers was a matter left up to individual districts. District officials affirmed this. For example, when asked about rewards and sanctions, some delineated between "the old way"—associated with No Child Left Behind (NCLB) and state-driven accountability measures—and "the new way"—often linked to the Every Student Succeeds Act (ESSA) and the "local control" afforded to individual districts. As one district administrator in Ohio expressed, "I feel like there's going to be more of that decision-making and that movement at the district level, versus, everything going back to the state . . . I just think there's going to be a lot more power [to make decisions about rewards and sanctions] at the district level under ESSA."

Importantly, this shift in power does not eliminate rewards and sanctions; instead it reconfigures decision-making for deploying rewards and sanctions at the district level. At the state-level, this has resulted in moves toward "soft power" that are manifested less in immediate intervention into district decision-making and more in the designation of macrolevel categories—for instance, "focus schools" or "schools of distinction"—that might provide incentives for districts that would like to obtain (or avoid) such a designation.

This "soft power" approach extends to the district. Across all three states, districts officials reported that they had moved from doling out rewards or sanctions based on assessments scores toward celebrating those that meet their goals for growth and providing support for those who are addressing a shortcoming. Still, categorizing schools as successful or in need of assistance are not neutral labels. Test scores and categories are made public, which means they still create pressure to perform well. Several district officials even mentioned the ways these scores are used in setting real estate prices in surrounding areas, confirming that even "soft" categories have an economic impact on communities. However, because they are not attached to a particular accountability mechanism as in previous reforms (Desimone, 2013; Cohen & Mehta, 2017), they are understood less as a traditional "carrot and stick" and more as a friendly nudge.

Along these lines, our data show a move away from using assessments as an explicit factor in evaluating teachers and administrators. Districts report more leeway in leveraging scores to target areas for improvement rather than punishment. When asked about local-level penalties for underperforming on an assessment, one urban Ohio district official said, "I wouldn't call them penalties—it would be more like, 'Okay, you have data showing this, then you need to be on an improvement plan . . . and see how you're going to make the changes you need to make." A suburban Ohio district administrator, likewise, took issue with sanctioning teachers for poor test scores. "Your goal is not to get rid of someone," the official said, "your goal is to make them better." In Texas, an urban district official described such an approach as a "longitudinal perspective"—that is, one that looks at teacher and school performance over time rather than punishing and rewarding based on fragmentary snapshots. With that said, even these "bottom-up" or "growth-based" approaches are not entirely disentangled from systems of rewards or sanctions. The same Texas official who spoke of longitudinal perspectives on accountability also talked about

their district's practice of holding "monthly recognitions," where trophies are distributed to schools showing growth in targeted areas.

Trend #5

If The World Were Perfect, It Wouldn't Be: **Despite Resource and Other Obstacles, The Standards Are Generally Accepted by Most Teachers, Districts, and States**

Despite instability in the policy environment, we found little resistance to standards from educators, even in the backdrop of the political backlash to Common Core state standards in some of our focal states. While previous research has documented both strong resistance and enthusiastic acceptance of standards (Cohen & Hill, 2000; Spillane, Reiser, & Reimer, 2002), we found little outright resistance or enthusiasm, but rather acceptance of standards as a useful tool, with an acknowledgement of challenges that accompany successful implementation. Our survey results show modest authority (buy-in) scores, for state content standards for teachers across teacher types, grades, and geographic areas. Scores ranged from 2.30 to 2.78 (there were no neutral categories, teachers had to either strongly agree/agree or strongly disagree/disagree). Most teachers reported similar scores, with the exception that teachers of students with disabilities (compared with general education and ELL teachers) and rural teachers (compared with suburban teachers) who reported significantly lower buy-in to the standards. While SWD teachers and rural teachers reported significantly lower buy-in to the standards overall, the size of these differences were quite small, and in Kentucky, rural teachers did not have lower buy-in, as they did in Ohio and Texas.

Our case study findings show even stronger authority for the standards: across all of our case study sites, state and district officials accept the standards and see them as important. We found no pushback against their existence and use except occasional anecdotes of isolated parent resistance in rural districts. Across all three states, this significant emphasis on local control seems to be connected to a relatively high level of authority of the standards. In other words, as district officials

reported greater autonomy in making decisions about how the standards would be implemented, they also reported increased buy-in. While several district officials in the three states suggested the need for minor revisions to the standards (e.g., reducing the number of standards or moving particular content from one grade to another), there was broad acceptance that the standards themselves are appropriate tools for driving student learning.

Trend #6: *You've Got To Be Careful If You Don't Know Where You Are Going, You Might Not Get There*: National Organizations Are Providing Direction to States & Districts in Standards Policies for English Language Learners

While the national pendulum has shifted toward local control for general education, standards implementation efforts for ELLs have become more centralized. Before NCLB, there was a lack of consensus about how to identify and reclassify ELLs (Olsen, 1989) and even into the 1990s most states lacked common identification and reclassification procedures (Cardoza, 1986) and there was considerable variation in definitions and procedures even within districts in the same state (O'Malley & Valdex Pierce, 1994). Our data indicate that since the passage of NCLB the ELL terrain has greatly changed, with states and national consortia taking more active roles in developing identification and reclassification guidelines and procedures local districts are expected to adhere to. Two national consortia have taken the lead; WIDA and the English Language Proficiency Assessment for the 21st Century (ELPA21) play a major role in helping states and districts provide the specificity, consistency, and authority needed for high-quality implementation of their state's standards. WIDA provides these directions to 37 states and Washington DC with ELPA21 providing these directions to 7 states, leaving only 6 states developing their own policies without support from either consortium.

Both organizations provide detailed guidelines, materials, and resources for teachers. WIDA provides guidelines for entry and exit criteria for the English Language Proficiency (ELP) program.

It also supports states in increasing the consistency between standards and assessments for ELP and general education, with the goal of ensuring that (a) instruction offered by ELL teachers prepares students to engage in the language demands of mainstream classrooms and (b) ELLs who score as fully English proficient on ELP assessments are ready to effectively engage with grade-level content. ELPA21, a newer consortium started in 2012, is developing a similar infrastructure of support for its partner states. While both consortia offer guidelines to their partner states in how to use the tools that they have developed for them, ultimately states decide which guidelines to adopt.

The reputation and stability of WIDA, which has existed since 2002, and the flexibility inherent in the use of its supports, may contribute to its authority with educators. This is especially true in Kentucky, which adopted the standards in 2006 before CCSS. As one Kentucky district official described, "they've been using WIDA standards here forever." The stability of their partnership has led to what one state official described as a family relationship. Local districts find the infrastructure provided by WIDA appealing in that it offers clearer guidelines related to how to support ELLs, as reflected by a Kentucky district leader's statement describing WIDA's measurement approach: "They have something called the 'Can Do Descriptors' . . . [which] say if the student is scoring at this level across listening, reading, writing, and speaking they can do these things . . . so that we see they give them some degree of predictability in terms of language acquisition growth." Because ELPA21 is still relatively new in comparison, the same stability and authority has not been fully developed.

Thus WIDA and ELPA21 are playing an increasingly prominent role in shaping policies related to standards implementation for ELs. This includes the development of ELP standards that align with the CCR standards, the development of ELP assessments that align with the CCR assessments and guidance related to cut-off scores for classification and reclassification of ELs. Districts indicate that they are "following WIDA" in shaping how they address standards for EL,

thus WIDA and ELPA21 might be thought of as de facto standards policies, given that state policy does not directly address how EL teachers could or should address instructional decisions for ELLs related to standards. In other words, WIDA and ELPA21 are the primary mechanisms for increasing the specificity, authority and consistency of EL policies as they relate standards implementation.

This centralization does not seem controversial. In fact, many district officials told us that they hope to receive more specific guidelines from the national consortia in how best to meet the needs of their ELLs. Perhaps this work has avoided the controversy of other efforts at centralization because support is designed as guidelines rather than mandates and it is led by private, not governmental, organizations.

Trend #7

***I Wish I Had An Answer To That Because I'm Tired of Answering That Question*: For Special Education, Tensions Between Individualization & Standardization Persist**

We focused only on students with disabilities who take the general state assessment. Compared with teachers of general education students, teachers of students with disabilities do not fully buy into the idea of standards as appropriate for their students, estimating that more than half will fail to reach grade-level standards (Edgerton, Fuchs, & Fuchs, forthcoming). Further, special education teachers reported low levels of specificity in terms of the guidance provided to them for how to help their students meet the standards.

While the authority for standards and specificity of implementation guidance for SWDs may be low, study respondents reported increased forms of power in implementing standards for SWDs. In interviews, state and district officials across the three states recognized that their efforts at guidance were primarily compliance related, helping teachers and districts avoid federal sanctions. Immense paperwork, substantial interaction with parents, and fear of being sued drove their day-to-

day standards implementation efforts. As one state official in Ohio described, SWD teachers "come out really excited . . . they've been dreaming of the first day of school . . . and they can get down in probably two months, of 'I can't believe I have this much paperwork to do, I can't believe IEPs (Individualized Education Programs) take this long to write."

The enduring paradigm mismatch between standardization and individualization emerged as a core explanation for the higher levels of power and lower levels of authority and specificity that we found for teachers of SWDs. Special education relies on the individualized education program, which is in tension with having the same standards for all students, and there are very few approaches to address this pervasive gap (Fuchs et al., 2015; Voltz & Fore III, 2006). During prior periods of standards-based reform, SWDs were often caught between principles of standardization—holding every student to the same standards—and guarantees of differentiation--receiving individualized instruction to meet a student's unique learning needs. This core tension persists across our data: study respondents wonder whether it is fair for students with IEPs to receive substantively different instruction yet be held to the same standards and take the same assessment. There were questions about whether both the *goals* of instruction should be differentiated (i.e., standards) as well as the *instruction itself* (i.e., pedagogy).

Our district respondents came down on different sides of these issues. One Texas SWD district specialist commented, "If they were performing on grade level and able to access the STAAR [the State of Texas Assessment of Academic Readiness] as it is, why would they need Special Education?" In Ohio, by contrast, the extended set of state standards for SWDs made them more appropriate for SWDs in the minds of district officials there, as these standards gave "those teachers roadmaps so that way they know the expectations… It gives them that continuum for every single standard, it's broken down." Kentucky district officials, through their focus on co-teaching and inclusion, supported the notion that the grade-level standards are indeed appropriate for SWDs.

Their "pretty high rate of inclusion in the general education [setting], 80% or more of the day" is an indication that the same set of standards apply to all levels.

While we found teachers of SWDs struggling with the same issues they have been for decades (Edgerton, Fuchs, & Fuchs, forthcoming; McDonnell, McLaughlin, & Morison, 1997), we identified several districts taking a pro-active approach, employing new models of co-teaching and co-PD, or using evidence-based approaches to support teachers in designing effective instruction for SWDs. Teachers of SWDs reported wanting more help from districts with technologies and curricula, but they also reported using and finding helpful the materials their districts provided. We did not find evidence that would bolster popular narratives that SWD teachers loathe the standards and associated tests; instead, we found moderate levels of support and a desire for more guidance on how to improve standards implementation for SWDs.

It's Tough To Make Predictions. Especially About the Future:

A Bright Future for the New Wave of Standards?

Our findings about the new wave of college- and career-ready standards in three states offer important implications for research, policy, and practice, suggesting a path forward for educators and policy-makers interested in creating more equitable outcomes for all learners. One of our most significant findings is the move toward local control in our focal states, with districts taking on a more prominent role in building the specificity, authority, and consistency of standards policy. Though calls for 'local control' might not seem particularly new, especially given the pendulum swings of standards reform, we found that the rhetoric of 'local control' was paired with concrete actions at the state and district levels to generate greater buy-in from key stakeholders.

Our results highlight the importance of *authority* in the current push for local control—understanding and believing in the standards as a useful and productive mechanism for improving teaching and learning. Given that previous waves of standards reform have been critiqued for top-

down, punitive approaches, it is notable that authority is being reinforced at the policy level to include educators' voices (especially ELL and SWD teachers). Policymakers may do well to make building such authority an explicit part of standards policy, through mechanisms such as townhalls, shared governance, feedback loops, and leadership efforts.

Though efforts toward increasing stakeholder buy in seem to be an effective means of generating support for standards reform, we found that an emphasis on local control includes challenges. One of the central challenges we identified involved districts struggling with capacity to provide the support and guidance that states once provided. Considering prior critiques of how standards reforms have disadvantaged under-resourced schools, we wonder how the recent push toward increased local control, particularly for under-resourced districts, will exacerbate existing inequities. Another challenge involved the age-old tensions surrounding special education in a standards system that simultaneously holds all students accountable to the same assessment and standards, while providing individual learning goals and objectives for SWDs (as well as ELLs and other students). While our study emphasizes the need for targeting resources to help teachers and school leaders address these pervasive tensions, we found that districts are attempting to address this challenge locally through investments in teachers (e.g., hiring additional teachers specializing in teaching SWDs and preparing all general education teachers for instruction of SWDs and ELLs) and selecting evidence-based curricula (e.g., those emphasizing differentiation). While we recognize the significant efforts that local networks of educators and district officials are making to support standards implementation on-the-ground in our three focal states, we remain cautious about whether these networks can be sustained over time, particularly as issues of resources and capacity strain the budgets and human infrastructure needed for successful policy implementation in the long term.

Despite the challenges reported by states and districts, they appear to be developing innovations to build authority, specificity, consistency, and power into their standards implementation efforts. One promising policy approach, for example, may be inter-district collaboration—to share best practices, help build local leadership capacity, and capitalize and leverage each other's resources and expertise—rather than each district individually taking on the tasks of developing aligned curricula and PD. More generally, we found that district efforts to implement substantive amounts of PD responsive to local contexts might serve as one lever for creating generative local policy environments for standards reform. Another promising finding relates to the increased role that two national consortia—WIDA and ELPA12—play in providing guidance to states and districts across the country to support the specificity, consistency, and authority of policies designed to assist ELL students in meeting state standards. This centralization of resources through national ELL consortia might suggest that more centralized resources could help lift the burden on districts (particularly under-resourced ones), and simultaneously provide productive resources and guidelines that can be locally adapted.

Alongside our findings about the increased role of authority in standards implementation in Texas, Ohio, and Kentucky, we found that the nature of state power has shifted, with previously stronger state power, or accountability, shifting to local "softer" power. This 'soft power' approach focuses on providing rewards and incentives for improvement rather than punitive sanctions, though district officials in all three states recognized that the possibility of sanctions still loomed in the background. In documenting the relationships between authority and power, we consider how stakeholder engagement, a mechanism for authority, may play a role in a 'soft power' approach to implementing state CCR standards. While we recognize the need for further research in exploring the interactions of the policy attributes in state and district implementation policies, our initial analyses of the interactions between the policy attributes (specificity, authority, power, consistency,

stability) suggest that this round of standards reform may meet with more success than previous attempts.

Funding

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant #R305C150007. The opinions expressed are those of the authors.

References

Anglum, C., Desimone, L.M., & *Hill, K. (forthcoming). Integrating Computer-based Curricula in the Classroom: Lessons from a Blended Learning Intervention. *Teachers College Record.*

Atkinson, P. Coffey, A., & Delamont. S. (2003). *Key themes in qualitative research: continuities and change*. Walnut Creek, CA: AltaMira Press.

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

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Education Research Journal, 42*(2), 231-268

Cardoza, D. (1986). The identification and reclassification of limited English proficient students: A study of entry and exit classification criteria. *NABE Journal,* 11, 21-45.

Carmichael, S. B., Martino, G., Porter-Magee, K., & Wilson, W. S. (2010). *The state of the state standards—and the Common Core—in 2010*. Washington, DC: Thomas B. Fordham Institute. Retrieved from https://files.eric.ed.gov/fulltext/ED516607.pdf.

Clune, W. H. (1993). Systemic educational policy: A conceptual framework. In S. H. Fuhrman (Ed.), *Designing coherent educational policy* (pp. 125–140). San Francisco: Jossey-Bass.

Coffey, A., Atkinson, P. (1996). *Making Sense of Qualitative Data Analysis: Complementary Strategies*. Thousand Oaks CA: Sage.

Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294-343.

Cohen, D., & Mehta, J. (2017). Why reform sometimes succeeds: Understanding the conditions that produce reforms that last. *American Educational Research Journal, 54*(5), 644–690.

Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record*, *106*(6), 1047-1085.

Desimone, L. (2002). How can comprehensive school reform models be successfully implemented?. *Review of Educational Research*, *72*(3), 433-479.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199.

Desimone, L. (2013). Teacher and administrator responses to standards-based reform. *Teachers College Record, 115*(8), 1-53.

Desimone, L. M. (2013). Reform before NCLB. *Phi Delta Kappan*, *94*(8), 59-61.

Desimone, L.M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory Into Practice*, *56*(1), pp. 3-12.

Desimone, L. & Garet, M. (2015). Best practices in teacher's professional development in the United States. *Psychology, Society, & Education*. 7(3), 252-263.

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

Edgerton, A. K., & Desimone, L. M. (2018). Mind the gaps: Differences in how teachers, principals, and districts experience college- and career-readiness policies. Working paper. Philadelphia, PA: University of Pennsylvania.

Edgerton, A. K., Fuchs, D., & Fuchs, L. (forthcoming). New standards and old divides: Policy attitudes about college- and career-readiness standards for students with disabilities. *Teachers College Record.*

Fuchs, L. S., Fuchs, D., Compton, D. L., Wehby, J., Schumacher, R. F., Gersten, R., & Jordan, N. C. (2015). Inclusion versus specialized intervention for very-low-performing students: What does access mean in an era of academic challenge? *Exceptional Children, 81*(2), 134-157.

Grissom, J.A., & Herrington, C.D. (2012). Struggling for coherence and control: The new politics of intergovernmental relations in education. *Educational Policy*, *26*(1), 3-14.

Grissom, J.A., & Strunk, K. (2010). Do strong unions shape district policies? Collective bargaining, teacher contract restrictiveness, and the political power of teachers' unions. *Educational Evaluation and Policy Analysis*, *32*(3), 389-406.

Hochberg, E., & Desimone, L.M. (2010). Professional development in the accountability context: Building capacity to achieve standards, *Educational Psychologist*, *45*: 2, 89—106.

Marsh, J.A., & Wohlstetter, P. (2013) Recent trends in intergovernmental relations: The resurgence of local actors in education policy. *Educational Researcher*, *42*, 276-283.

McDonnell, L. M., McLaughlin, M. J., & Morison, P. (Eds.). (1997). *Educating one and all: Students with disabilities and standards-based reform* (Report of the Committee on Goals 2000 and the Inclusion of Students with Disabilities). Washington, DC: National Research Council. http://www.nap.edu/catalog/5788.html.

Mehta, J. (2013). *The allure of order: High hopes, dashed expectations and the troubled quest to remake American schooling.* New York, NY: Oxford University Press.

Olsen, R. (1989). A survey of limited English proficient student enrollment and identification criteria. *TESOL Quarterly,* 23, 469-488.

O'Malley, J. & Valdez Pierce, L. (1994). State assessment policies, practices, and language minority students. *Educational Assessment,* 2, 213-255.

Pak, K., & Desimone, L. (forthcoming). How Do States Implement College and Career Readiness Standards? A State Distributed Leadership Analysis. *Educational Administration Quarterly*.

Polikoff, M. S. (2012). The association of state policy attributes with teachers' instructional alignment. *Educational Evaluation and Policy Analysis, 34*(3), 278–294.

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

Porter, A. C., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1988). Content determinants in elementary school mathematics. In D.A. Grouws & T.J. Cooney (Eds.), *Perspectives on research on effective mathematical teaching* (pp. 96–113). Hillsdale, NJ: Erlbaum.

Ravitch, S. M., & Carl, N.M. (2016). *Qualitative research: Bridging the conceptual, theoretical, and methodological.* Thousand Oaks, CA: SAGE Publications.

Ryan, G. W. & Bernard, H. R. (2003). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.), *Collecting and interpreting qualitative materials* (2nd ed.) (pp 259-309). Thousand Oaks, CA: Sage.

Spillane, J. P. (2009). *Standards deviation: How schools misunderstand education policy.* Cambridge, MA: Harvard University Press

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing on implementation research. *Review of Educational Research*, *72*(3), 387-431.

Tyack, D. F. (1974). *The one best system: A history of American urban education.* Cambridge, MA: Harvard University Press.

Voltz, D. L., & Fore III, C. (2006). Urban special education in the context of standards-based reform. *Remedial and Special Education*, *27*(6), 329–336.

Wong, V.C, Wing, C., Martin, D., & Krishnamachari, A. (2017). Did states use implementation discretion to reduce the stringency of NCLB? Evidence from a database of state regulations. *Educational Researcher*, 47(1), 9-33.

Table 1. Description of Survey and Interview Respondents

	# District Survey Respondents (response rate)	# Principal Survey Respondents (response rate)	# Teacher Survey Respondents (response rate)	# State Interview Respondents	# District Interview Respondents
Kentucky	N/A	179 (50.6%)	740 (41.6%)	13	19
Ohio	42 (79.2%)	111 (60%)	417 (63.8%)	13	20
Texas	42 (85.7%)	149 (70.6%)	603 (55.4%)	10	15
Total	84	439	1760	36	54

Table 2. Characteristics of Interview-Study Districts

District	Total Population	SWD	ELL
KY urban	39,952	11%	9%
KY suburban	20,046	12%	5%
KY rural	1,077	16%	5%
OH urban	21,708	19%	5%
OH suburban	10,076	14%	3%
OH rural	561	16%	5%
TX urban	43,063	10%	22%
TX suburban	52,801	10%	13%
TX rural	19,500	9%	4%

Sources: National Center for Education Statistics; State Education Agency District Report Cards; United States Census Bureau Small Area Income and Poverty Estimates

Table 3
Cross-State survey comparison of teacher-reported standards-based professional development received in annual hours

	Texas				Ohio				Kentucky			
Type of Professional Development	N	Mean	Min	Max	N	Mean	Min	Max	N	Mean	Min	Max
Workshops, conferences, institutes, or seminars	556	31.80 (2.47)	0	185	394	31.80 (2.47)	0	225	351	19.10 (2.32)	1	84
Professional learning communities	556	34.43 (4.23)	0	1000	393	34.43 (4.23)	0	1000	375	28.96 (4.50)	1	200
Formal coaching or mentoring	551	9.73 (1.50)	0	300	392	9.73 (1.50)	0	200	375	12.53 (1.50)	1	180
Multiweek course	546	4.32 (0.92)	0	500	386	4.32 (0.92)	0	500	375	6.53 (0.92)	1	240
Total Hours (1 = 1–10, 2 = 12–30, 3 = 21–40, 4 = 41–80, 5 = 80+)	601	1.74 (0.06)	1	5	416	1.43 (0.08)	1	5	404	1.31 (0.05)	1	5

Standard errors in parentheses after survey weights were used.
* $p < 0.05$, ** $p < 0.01$

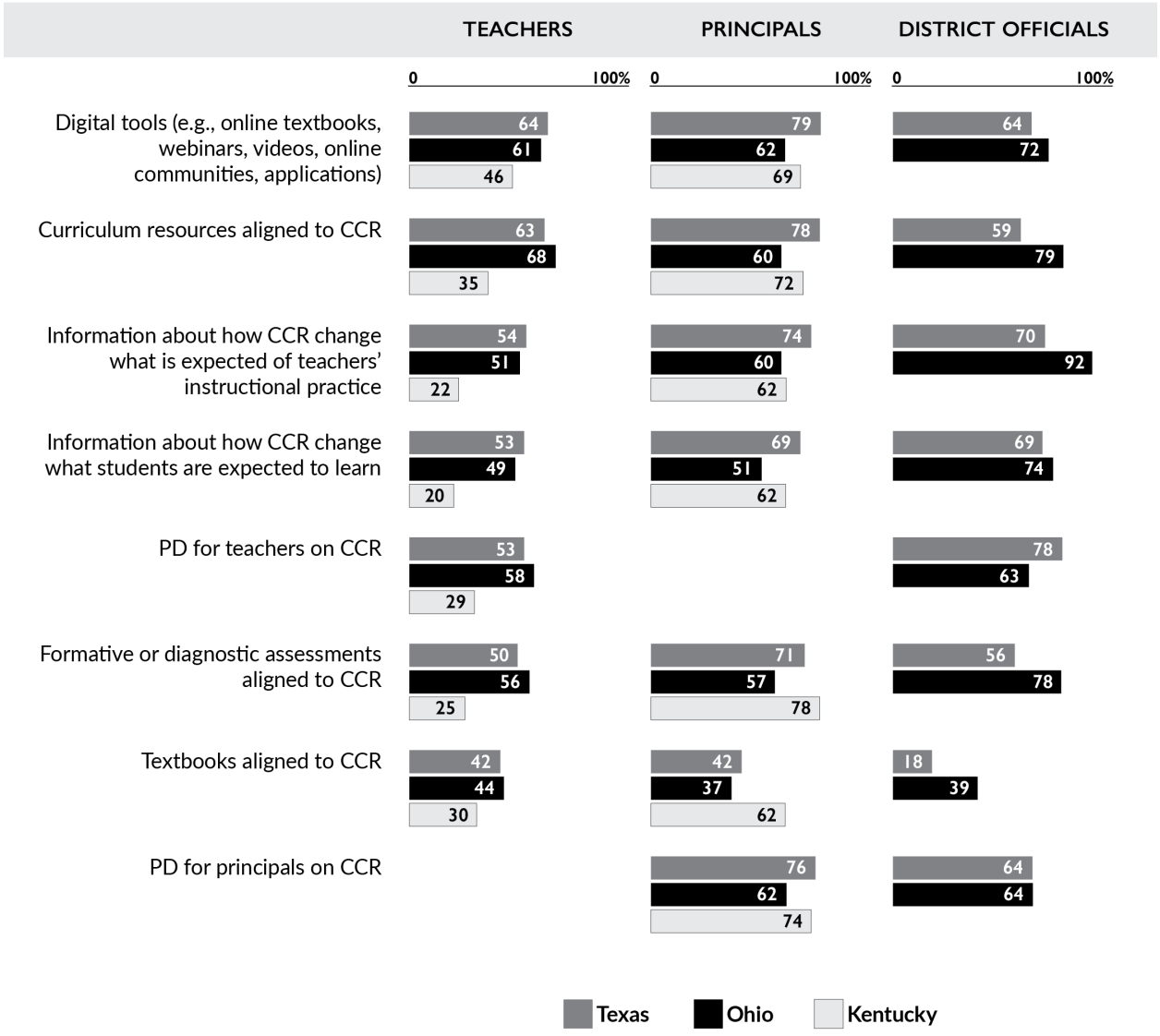
Table 4
Within-State survey comparisons of teacher, principal, and district survey responses

Texas	T N	Mean	P N	Mean	D N	Mean	F Test	Sig.
Specificity	585	3.14	158	2.94	41	3.35	1.54	
		(0.07)		(0.09)		(0.11)		
Consistency	564	2.82	166	2.83	42	2.75	0.13	
		(0.04)		(0.06)		(0.13)		
Authority	583	2.56	178	2.95	42	2.73	8.20**	T v. P
		(0.05)		(0.07)		(0.11)		
Power	586	2.68	153	2.35	42	2.34	3.47*	T v. P
		(0.07)		(0.10)		(0.16)		
Stability	579	2.51	153	2.74	41	2.92	2.36	
		(0.07)		(0.09)		(0.17)		
Ohio	T N	Mean	P N	Mean	D N	Mean	F Test	Sig.
Specificity	405	2.38	110	2.60	42	3.03	1.69	
		(0.13)		(0.12)		(0.18)		
Consistency	379	2.71	108	2.83	42	2.74	1.13	
		(0.04)		(0.06)		(0.08)		
Authority	402	2.30	109	2.90	42	2.50	23.58**	T v. P, T v. D
		(0.04)		(0.08)		(0.13)		
Power	405	2.50	110	2.38	42	1.96	4.08*	T v. D
		(0.06)		(0.12)		(0.11)		
Stability	398	2.44	109	2.83	42	2.28	5.36**	T v. P, P v. D.

		(0.06)		(0.11)	(0.20)	
Kentucky	T N	Mean	P N	Mean		T Test
Specificity	436	2.75	158	2.84		1.14
		(0.04)		(0.07)		
Consistency	390	2.91	166	2.52		6.92**
		(0.03)		(0.05)		
Authority	430	2.78	159	3.18		6.52**
		(0.03)		(0.06)		
Power	428	2.56	153	1.96		8.66**
		(0.03)		(0.08)		
Stability	400	2.29	N/A	N/A		
		(0.06)		N/A		

Standard errors in parentheses after survey weights were used. Responses ranged from 1-4, where 1 = strongly disagree, 2 = somewhat disagree, 3 = somewhat agree, and 4 = strongly agree for relevant items in the policy attribute construct.
* $p < 0.05$, ** $p < 0.01$

Figure 1
Within- and cross-state survey comparison of resources desired by teachers, principals, and district officials*



CCR=College and Career Ready Standards and PD=Professional Development
*Due to our partnership arrangements with Kentucky, we were not able to administer a district survey.