

The Current Status and Role of Standards Based Reform in the States

By

Diane Massell, PhD
University of Michigan

with research assistance from

Mark Duffy
Stephanie Levin
Caitlin Scott

*A draft report prepared for the National Research Council Workshop Series on
Standards Based Reform: January 17-18, 2008*

This report seeks to inform the workshop proceedings of the Hunt Institute on the role of standards reform in K-12 education by asking four central questions:

- (1) What are the major roles that standards play in state K-12 education policy and practice?
- (2) What are the major strengths and weaknesses of K-12 state standards-based reform efforts with respect to achieving efficiency, equity, and quality? What are states doing to achieve these goals?
- (3) How and to what degree are the strengths and weaknesses of reform efforts related to the standards themselves?
- (4) How and to what degree are the strengths and weaknesses of reform efforts related to having unique state standards?

To answer these questions, the research team conducted a series of interviews in a sample of five states. The purpose of these interviews was to elicit broad themes and insights to inform the workshop discussions, and illustrate findings and conclusions from the body of extant research on standards-based reform. The latter draws heavily on research conducted or synthesized by the Consortium for Policy Research in Education, as well as other literatures.

The five states—California, Florida, Massachusetts, North Dakota and Texas—represent different regions of the country and different histories and experiences with standards-based reform. California, for example, was an early pioneer in the standards movement, engaging in ambitious efforts to set challenging state guidance policies as early as the mid-1980s. North Dakota was a relative latecomer to these reforms, initiating their efforts after federal law required standards in 1994.

Between November and December, 2007, researchers conducted telephone interviews with 21 policy elites, ranging from 3-6 per state. Researchers selected policymakers who were engaged in education reform, and included officials or education aides from the governors' offices, state boards of education, state legislatures, and state education agencies.

Role of Standards in K-12 Education Policy and Practice

As the modern standards-based reform movement enters its second decade, our findings and other research suggests that standards remains a central framework guiding state education policy and practice. Indeed, policymakers in the four states with the longest history of involvement in these initiatives—California, Florida, Massachusetts, and Texas—were in universal agreement on this point. For example, when asked what role standards played in their current K-12 strategy, one Texas lawmaker replied, “Gigantic! Everything is based on the standards.” Florida officials believed that standards had become even more central over time. California and Massachusetts officials used words like “integral” “profoundly” and “predominant” to describe the importance of standards to their policy initiatives. They also concurred that standards had a strong, if variable, impact on local practices. Such variability was partially attributed to a range of capacity issues, including insufficient knowledge and skill or fiscal resources, and weak institutional capacity of the state and some local districts to support teachers and administrators, among other things (see below).

Evidence of the centrality of standards reforms can be found in the growing list of policies and resources to which they have been aligned. Since 1994, federal law has required that all states set challenging standards aligned to assessments and

accountability. In these states, such alignment had been established or was underway prior to federal mandates. In addition, standards were also aligned or in the process of being linked to many other policies, such as state-provided or funded professional development, curriculum frameworks, high school course descriptions, model lessons, test specifications and/or released test items, teacher preparation programs, teacher certification and licensure, funding or other supports for low- or even high-performing schools and districts, and teacher performance pay incentives. The three states that adopted textbook and materials statewide—California, Florida and Texas—also used standards to guide their selection.

To be sure, standards reforms continue to galvanize political disputes, particularly over testing, performance standards, and accountability. In several of the states reviewed here, teachers and teacher organizations voiced criticism over the emphasis on state test scores to determine their accountability status, as well as the high volume of state testing. States like Massachusetts and North Dakota expanded their testing programs substantially after the passage of No Child Left Behind (NCLB), which requires annual testing in grades 3-8 and once in high school. Teachers expressed frustration over the challenge of addressing numerous state content standards, determining which standards were most essential, and the lack of flexibility and creativity in the standards. Not surprisingly, states that were considering or had passed initiatives to link teacher pay incentives to improvements in student achievement stirred the opposition of teacher organizations. Finally, teachers and others were concerned about the narrowing or constraining influence of testing and accountability on curriculum and instruction—a view supported by some evidence, though not always shared by the policymakers we interviewed (McMurrer 2007; see also Stullich et al. 2006).

In several states, business organizations advocated for more rigorous performance standards, especially at the high school level, and at least in one state, to more stringent application of NCLB accountability sanctions like school closure for chronically underperforming schools. Teacher, superintendents and community groups were more likely to object to such sanctions, and express concern over the potentially negative effect that higher performance standards would have on students and schools. For example, opinion research showed that while teachers support high school exit exams, most

support a basic skills test (62%) rather than a more challenging exam (24%) (Reality Check 2006). Educators and policymakers in many states dispute the validity and accuracy of accountability measures, especially under No Child Left Behind, and three of the states here—California, Florida and Texas—maintained their own, separate accountability programs on top of NCLB.

Finally, teachers groups, professional associations and other special interest groups voiced disagreements with particular aspects of state content standards, issues that emerged as states undertook efforts to update their standards. Four of the five states—Florida, North Dakota, Massachusetts and Texas—were revising their standards, representing some of the 37 states that have adopted or revised academic standards since NCLB passed in 2001. Although several California policymakers thought the state needed to update their standards, they noted that a long-standing stalemate between the governor and the legislature over control of standards prevented any progress in this regard (see Fordham Foundation 2006).

Many of the tensions that arose in the development of standards have been present since the beginning of the standards reform movement (Massell and Kirst 1994.) For example, debates emerged over how specific or broad and flexible standards should be; whether standards should promote particular instructional approaches and if so, which ones; how to present standards that sufficiently cover important content without being overly broad and unfeasible for teachers to implement in their practice. Certain disciplines, such as science and history, were also seen as more likely to generate political conflict.

However, despite these specific issues and concerns, educators and others have grown to support and value standards-based reform. Other studies have shown that teachers appreciate the clear identification of what students should know and be able to do, and report using standards and assessments to inform their own practice. While teachers are less likely to view state test scores as an accurate reflection of student achievement than administrators, they do find it reasonable to hold students and educators accountable for their performance and use this data to inform their practice. They perceive that standards reforms have increased curriculum coordination within their schools, improved the focus on student achievement, and increased rigor (Goertz and

Massell 2005; Massell et al. 2005; Hamilton et al 2005; Reality Check 2006; Herman 2007). Focus groups and telephone interviews also show that support for standards is broad among parents, teachers and administrators (Reality Check 2006).

Even though NCLB accountability became a lightning rod of dissent in many states, some believed that NCLB had in fact helped to legitimize accountability. For example, one Massachusetts policymaker said that NCLB “makes it apparent that accountability is here to stay” and noted that it had made the disaggregation of data and the reporting of subgroup performance more acceptable to various interest groups. Similarly, a Texas legislator commented that “NCLB was a big step forward,” and acknowledged that while they had a difficult time meeting some of the accountability criteria, “states have come to realize that the bottom line of NCLB is something that we all very much need.” Furthermore, fifteen of the 21 policy elites interviewed here stated that one of the most positive aspects of standards was that it had in fact brought various stakeholders together and developed a common understanding across diverse communities about academic content and performance expectations.

In contrast to the other states reviewed here, three out of the four North Dakota officials did not see standards reforms as central in state policy. They also thought that standards had only minimal to marginal effects on classroom practice. There are several possible explanations. Unlike the others, North Dakota’s initiatives did not stem from any significant internal political impetus or demand, but rather in response to the federal Improving America’s Schools Act of 1994 (IASA). As already noted, California was a forerunner of the current standards movement, developing a more expansive state role in establishing challenging instructional guidance in the 1980s. Massachusetts also began to entertain standards-based reform as a solution to a school finance lawsuit as early as 1988. At that time the Massachusetts Business Alliance for Education proposed coupling funding increases with a system of standards, assessments and accountability, ideas that eventually led to the 1993 Massachusetts Education Reform Act (McDermott 2004). Florida and Texas also assumed strong leadership over instructional guidance policies prior to 1994. While North Dakota officials report no significant opposition to standards reforms today, they also did not identify any significant interest groups who advocated for improving them, and the business community here was not seen as an instrumental

player in state education policy. Finally, while North Dakota’s accountability program relies on its standards-based assessments and performance measures as required under federal law, interviews indicate that alignment and connections to other policies were largely left to local districts to develop.

While one cannot draw any strong conclusions based on this small sample, our evidence suggests that while standards reform is ubiquitous on state policy agendas, the depth of commitment and focus on standards reform still remains variable across different states. But the preponderance of extant research does indicate that for the most part, standards reforms have become a legitimate and taken-for-granted feature of education policy regimes, even by their most vocal critics—something that many observers would not have predicted at the start of this reform movement.

Major Strengths of Standards Reform

The policymakers interviewed for this report were asked a series of questions about the impact of standards reforms on practice, and on learning opportunities, the quality of education, and resources in particular. Their responses suggest that they perceived more equitable learning opportunities for traditionally underserved populations as the most positive outcome and the most promising potential of standards-based reform and accountability (see similar findings from Center for Education Policy 2006). Indeed, when asked about the effects of standards on the quality of education in their state, most immediately began discussing equity. At the same time, they saw improving learning opportunities and continuing to make achievement gains as some of the most difficult challenges they faced.

Specifically, they believed that five types of equity had improved under the regime of standards-based reform:

(1) Greater awareness and attention to the academic performance of underserved students. Standards reforms, and most especially the disaggregation of student achievement data as required by NCLB, raised public and professional attention about the performance of students from different populations, including minorities, limited English

proficient, low income, and special education subgroups. For instance, one North Dakota policymaker commented that because of NCLB disaggregation rules they are more aware of their failings with special education and native American students. These NCLB requirements prompted numerous in many state assessment and accountability programs, even in states like Texas that had served as a model for NCLB and had long disaggregated student achievement data. One need only recollect the outcry over the proposal of former US Secretary of Education, Terrell Bell, to publish a “Wall Chart” in 1984 reporting aggregate, state-level results on the SAT and ACT to realize how far we have come as a nation in our acceptance of comparisons and the value we place on reporting data about our students and schools.

(2) Higher expectations that all students will achieve to more rigorous standards. Standards provides educators with an existence proof that students can learn to higher standards if the system expects them to do so. Said one California lawmaker, “Standards got us past the discrimination of low expectations” for students in poverty, of different ethnicities, and of limited English abilities. Indeed, standards provided states with some political leverage for pressing higher expectations. For example, when people complained that their special education students would not reach 100% proficiency, one North Dakota official reported showing them data that other schools had made greater progress with these students. “You can look them in the eye and tell them they need to improve. It cuts the oppositional language off at the knees.”

(3) Reductions in the achievement gap. While acknowledging that much remains to be done, several policymakers noted that standards reforms had reduced the performance gap between students from different backgrounds, and between advantaged and disadvantaged schools. 2005 National Assessment of Educational Progress (NAEP) reports show progress in narrowing the achievement differences between subpopulations, especially at the elementary level, and in mathematics. Other analyses of state assessment data indicate that in the last few years, achievement gaps have narrowed in a majority of the states they reviewed (Center on Education Policy 2007; Education Trust 2006).

(4) A more uniform educational system within states. Most policymakers agreed that standards reforms had resulted in a more uniform system of education in their states, though they also recognized that implementation and practice varied. One California policymaker commented, “Standards have been, for many of those groups, the most important civil rights effort we have seen in many, many years,” by giving students from underserved populations more access to a common core curriculum. Another argued that standards-based assessments had helped to forge uniformity by calibrating expectations for student performance to a common metric.

(5) Instruction tailored to the needs of different learners. The press to educate all students to state standards was perceived to have contributed to more widespread use of or at least acceptance of differentiated instruction, where schools and teachers identify the particular learning needs of students and adjust their teaching accordingly.

If you are under standards and accountability and you have students reading at the second grade level and the college level in the same class, how can you meet the needs of all those kids? What we are finding is that some of our schools are moving toward ability grouping and able to have smaller classes for the struggling students while the high achieving students are not held back. It is a rethinking of what is equity and what is excellence so that we can move forward and deal with the needs of kids.

Florida has developed a statewide data system to enable sophisticated use of state testing information by schools and teachers, and many states are working towards or perceive the need to provide data systems that can similarly support their standards. Since the mid-1990s, there has been an explosion of district and school adoption of quarterly or other interim assessments to measure student progress towards standards and to provide formative feedback to teachers (Massell and Goertz 2002; Le Floch et al 2007). Many initiatives encourage and support the use of data for differentiated instruction, including comprehensive school reform designs like Success for All, America’s Choice, and the

Accelerated Schools Project, although evidence suggests that using data for such purposes is not always well understood or implemented in the classroom (Massell and Barnes 2006).

Major Challenges in Standards Reforms

While policymakers had positive views of the merits of standards, they also identified significant challenges. As noted above, improving equitable outcomes for all students was seen as one of the most difficult issues they faced, solutions to which are compounded by problems in the capacity of the educational system. Secondly, policymakers identified high school reform as a significant challenge that they were seeking to address in part through changes in standards. Third, these states faced a series of issues related to performance standards and accountability.

Improving Equitable Outcomes for all Students

As mentioned above, 2005 NAEP and 2006 state assessment results indicate reductions in the achievement gaps between different populations. But differences in performance do persist and are quite large. For example, only 13 percent of African American and 19 percent of Hispanic 4th graders scored at or above the proficient level on NAEP mathematics tests. By contrast, 47 percent of white students performed at or above these levels. On NAEP reading assessments, African American 17 year olds perform at the level of white 13 year olds. Similarly English language learners and students with disabilities also fall far behind on NAEP reading (NCES 2005). On state assessments, achievement differences between groups were also large—an average of 20 percentage points or more in the percentage proficient—and while gap reductions had been made in most states, many also experienced no change (Center on Education Policy 2007). Furthermore, the types of schools not meeting performance targets under NCLB accountability are more likely to be urban schools, and schools with large concentrations of poor or minority students. In 2004-05, for example, one third of high poverty schools did not make adequate yearly progress, compared to only four percent of low poverty

schools. One third of high minority schools and more than one-fifth of urban schools also repeatedly failed to meet performance targets (Le Floch et al. 2007).

Significant resource issues hindered or exacerbated states' ability to address equity concerns, as well as the implementation of standards reform more generally. Resource problems including weaknesses in the capacity of state education agencies and district organizations, funding adequacy and equity, the availability of qualified teachers particularly in urban or rural locations, adequate professional development and instructional designs that could improve the learning of all students, especially to the levels needed to meet accountability targets under NCLB.

(1) Capacity of States and Districts to Support Low-Performing Schools and Districts. Standards based reform, and NCLB in particular, have stretched the capacity of state education agencies and districts to assist low-performing schools. Over the last ten to twenty years, many state education agencies have lost substantial personnel and financial resources. For example, when the Massachusetts legislature passed its standards-based reform act in 1993, the state had 325 full-time staff members—less than the number of school districts in the state, and a drop from 990 employees thirteen years earlier (McDermott 2004). In addition to traditional responsibilities for monitoring compliance with rules and regulations, standards-based reforms required that these agencies develop or oversee the development of academic standards and their alignment to assessment, accountability and other policies. The difficulties of such alignment work and other technical issues such as measuring the progress of limited English proficient students are substantial, and still evolving (CCSSO 2007).

Under NCLB, the expectations for state agency staff have expanded even farther, including, but not limited to, a mandate for creating statewide systems of support to help all Title I schools and districts identified for improvement. The number of identified schools are large compared to the numbers typically identified by states prior to NCLB, and are expected to grow. In 2003-04, 25 percent of schools across the country did not make adequate yearly progress under NCLB, a proportion that hides considerable variation across states due to different performance standards and measures. But states like Florida and Alabama identified quite a large proportion -- approximately two-thirds

of their schools and more than 90 percent of districts. And more than half (27) of the states adopted a particular accountability design (a “delayed acceleration pattern”) which expects that two-thirds of achievement growth will occur after the first six years. It is more than likely that these states will experience an explosion of schools and districts not making adequate yearly progress (Le Floch et al. 2007). One Massachusetts policymaker interviewed for this report believed that the most difficult challenge they faced was improving state capacity to address the cascading number of schools being identified as under-performing.

Many states have struggled to determine how to design their statewide system of support and find the necessary staff and organizational resources to assist those in need (see Massell et al. 2006; Rennie Center 2005; CEP 2005). Some states have many nongovernmental organizations and strong intermediate education agencies that can assist them with these tasks. Others do not. But even when they do, managing the support can be a large and cumbersome task for state agency personnel.

While NCLB specifies certain features of the support system that states must provide, such as the creation of school support teams with distinguished educators and principals from successful schools, much is left to states in partnership with districts to determine. States and districts have undertaken many actions to assist low-performing schools—such as targeting resources, aligning curriculum and instruction, and analyzing achievement data. However, recent studies have shown that they often lack the capacity to provide assistance to all schools identified for improvement, including Title I schools not in the weakest performance categories, and non-Title I schools who are—particularly to provide the more intensive assistance of school support teams, school-based coaches, or principal mentors (Le Floch et al. 2007; see also Padilla et al., 2006 and CEP 2007). Capacity across districts also varies considerable, with larger districts tending to have more staff and resources than smaller districts and sometimes state agencies. Nevertheless, they too do not reach all identified schools with intensive resources (Padilla et al. 2006; for more on this topic, see the summary in Goertz 2007). Policymakers we interviewed for this report also identified other kinds of variation in district capacity as a factor in whether standards reforms were well implemented. This importance of district

leadership is supported by numerous studies of standards-based reform and implementation (Spillane; Barnes et al, forthcoming; Suppovitz 2007; Learning First).

(2) Adequate Funding and the Supply of Qualified Teachers The adequacy of school funding and the availability and equitable distribution of highly qualified teachers are two other resources challenges that states face in their efforts to improve learning opportunities. Adequate funding and funding equity was a major issue in California and North Dakota, as in other states around the nation. In California, a class action suit was filed in 2000 charging that the state had failed to provide students with equal access to instructional materials, qualified teachers, and safe and decent school facilities (*Williams et al vs. State of California*). The case was settled in 2004, resulting in the allocation of \$138 million in funding for standards-aligned instructional materials for low-performing schools in the first and second deciles of the state accountability system. Class size reduction policies in place in California as well as Florida were seen as exacerbating their ability to get highly qualified teachers into low-performing schools. Evidence suggests that experienced and highly credentialed teachers in California are more likely to teach in affluent, majority white schools (Education Trust-West 2005). Nationally, states have experienced difficulty in attracting and retaining qualified teachers in rural and urban settings.

(3) Instructional Knowledge. Finally, policymakers questioned whether teachers and administrators had the knowledge and skills they needed to continue to make progress on reducing the achievement gap, either because they did not have sufficient access to professional development, or because instructional designs geared toward improving the achievement of diverse learners were scarce (Commission on No Child Left Behind 2007).

High School Reform Issues

One concern across these states was the relative lack of improvement in high school performance and the perception that high schools have not always been a strong and sustained focus of state reform in general or standards reforms in particular.

Policymakers voiced particular alarm over graduates' lack of preparedness for careers as well as college remediation numbers. 2001 data from the National Center for Education Statistics, for example, showed that 40 percent of students at four-year postsecondary institutions, and 63 percent in two-year institutions, required some form of remediation (reported in Callan et al 2006).

Policymakers in several of these states were looking at various standards-related remedies, such as raising credit requirements, developing a stronger alignment between high school courses and college and workforce readiness standards, raising proficiency standards or developing a new high school exit exam¹. In fact, aligning high school standards to college and work readiness has been a prominent, recent theme in many national organizations and conferences (see, for example, the American Diploma Project of Achieve, Inc., or the proceedings of the National Education Summit on High School Reform in 2005). Texas adopted a new academic standards program called "4 x 4" that requires four years of English, math, science and social studies for graduation. They plan to connect this college readiness program to state end-of-course exams and content standards. Florida passed the Florida Career and Professional Education Act which requires every school district to enter into joint venture with the local workforce board to start a career institute, where according to one respondent, industry will establish the standards and determine the curriculum as well as the testing to meet industry standards. Students at these institutes could earn a high school diploma, postsecondary credits and national industry certification in career fields. These states are not alone in their efforts to target high school reform. By some counts, as many as 48 states now have or are developing policies that include higher graduation standards, new college and work readiness curricula, and more course credit requirements (Achieve 2007).

National business organizations and business communities in several of the states reviewed here want to see more rigorous performance expectations for high school students (US Chamber of Commerce 2006; Business Roundtable 2007). Many state exit exams have been geared to the end of middle school or the beginning of high school in

¹ Policymakers also addressed problems like low attendance and graduation rates with reforms not specifically associated with standards, such as ninth grade transition programs, summer school programs, and more.

part to prevent high drop out rates. In Massachusetts, for example, graduation proficiency standards were set low for this reason. But now the business alliance has called for adding more subjects to graduation exit requirements and raising proficiency standards for graduation (Massachusetts Business Alliance for Education 2006; McDermott 2004). These exams are not seen as challenging or motivating for more academically able or average students, and not adequately ensuring that high school graduates are prepared for college or work.

But many are concerned that raising high school standards would have negative impacts on disadvantaged populations. So, for example, while most the Texas officials we interviewed believed that their standards are too low, they feared that students might drop out or experience other harms as standards ratchet up over time. States are legally bound to assure that students have an opportunity to learn the materials covered by exit exams, and when states revise their academic or performance standards they have to ensure that this is accomplished. One California policymaker believed that high drop out rates had occurred because schools had not had sufficient time or resources to make the necessary adjustments. In its effort to demonstrate instructional validity, Florida was creating on-line professional development materials and assessments for teachers to determine if they know and understand the new standards.

Yet the political problems and questions remain. One quote from a Florida official aptly sums up the challenge of raising performance standards:

We have been going through a very public external review process of our assessment and accountability system in Florida over the last seven or eight months. The focus has been on expectations for the performance of students in high school. The concern is that they do not appear to perform as well as middle school and elementary school students. The question is whether or not we have set the expectations too high. Should we reduce them? Superintendents tell us that things are too high and not attainable. You tend to hear from business that they are not high enough. You hear that students leaving high school are not well prepared for the work world, nor are they prepared for post-secondary. You hear this from post-secondary people. Teachers tend to have a mixed reaction to the

standards...The press is like a teeter-totter. If the state is being attacked because the standards are too high then they focus on that. If the counter is that students are not well enough prepared for post-secondary then they focus on that.

Accountability Design Challenges

Of course, states need to determine appropriate performance standards not just for their high school assessments, but for their accountability determinations for schools and districts more broadly. While NCLB requires more uniformity in state assessment and accountability provisions than its predecessor, it still gives states the flexibility to determine their own performance standards, proficiency targets, ways of calculating progress, subgroup sizes, and more. As states made these decisions and negotiated their performance standards, they arrived at very different solutions. While research methodologies for comparing state performance standards to NAEP are problematic in different ways, the preponderance of evidence across different studies supports the view that state performance standards vary considerably (NCES 2007, Kingsbury et al 2003). Furthermore, when compared against NAEP, the overwhelming majority of states' performance standards are much lower. So, for example, while 88 percent of fourth graders in Tennessee scored at the proficient level or above on state reading tests, only 27 percent performed at the proficient level or above on NAEP. In California, high school standards are lower than NAEP. Massachusetts performance standards for different grade levels are consistently high when compared to NAEP.

Such differences have consequences for the numbers of schools which make annual yearly progress. States with high proficiency standards in fourth and eighth grade reading identified 30 percent of their schools as not meeting AYP in 2003-04. By comparison, states with lower standards only identified 16 percent of their schools (Study of State Implementation of Accountability and Teacher Quality Under NCLB, Interim Report 200x). As this suggests, states do not agree on what performance levels should be set, a variation that represents different assumptions about fairness and best practice, as well as different political and social considerations. For example, some are concerned that large numbers of failing schools reduce the power of the accountability system to motivate and focus improvement.

Many states have also questioned the fairness and accuracy of NCLB accountability. One key concern, and one that many proposals for NCLB reauthorization address, is NCLB's use of a "status bar" model of progress. NCLB specifies that states set a single performance standard for all schools and subgroups of students in reading and math, and progress is measured simply by whether schools meet these standards or not at any one point in time. States and others have argued that schools should be given credit for making progress towards achievement goals. One Texas policymaker noted that she, too, had come to the conclusion that "student growth is more important than how a student performs compared to other students in the school and the state." So called value-added accountability designs also provides a potential way for policymakers to measure teacher performance based on student achievement. While not uncontroversial, value-added growth models have become an increasingly popular consideration for states, and in 2005 the US Department of Education announced approval for up to ten state pilots.

In addition, some have also questioned the fairness of using subgroup population calculations to determine schools' accountability status, since more homogeneous schools are more likely to meet their performance targets than highly diverse schools (Sunderman et al. 2005). Measuring the progress of English language learners is also problematic because they tend to be a highly mobile population which makes it difficult to measure progress from year to year. For example, in Miami-Dade, Florida, approximately 25,000 English language learners enter or leave the district each year (Commission on NCLB 2007). Furthermore, students who become proficient are no longer included in the tested subgroup, thus underestimating progress (Zabal 2007).

For these and other reasons, at least eleven states, including California, Florida and Texas, have maintained their own, separate state accountability program alongside NCLB. These separate state systems vary in several ways from NCLB. Some, like California and Florida, reward growth. Some do not include subgroup calculations. As a result, these state programs tend to identify fewer schools as low-performing (Goertz 2007; Goertz 2005). In 2004, 1,626 California schools were identified as underperforming in NCLB calculations; by contrast, 71 percent of these schools met improvement targets under the state system (Sunderman et al, 2005). While states have

strong arguments for their own designs, dual accountability systems have generated substantial confusion among educators and the public alike. One Texas official observed that “two accountability systems...is a complication for everybody. It confuses people because you can be okay in one system and not in the other. How do schools figure out which one to focus on for.. standards?”

Unique or Common Standards?

In the past year, many have renewed the call to establish a voluntary set of common, national content and performance standards and more cross-state comparisons (see, for example, American Federation of Teachers 2006, Commission on No Child Left Behind 2007, the Education Trust, 2007). These proposals would pursue this goal in different ways. For example, the American Federation of Teachers supports funded explorations of “shared standards,” while the Commission on NCLB argues for the mandatory use or incorporation of model national standards based on NAEP frameworks. The Education Trust suggests incentives for adopting assessments and performance standards as well as sanctions for states with large discrepancies between proficiency rates on their own assessments and NAEP results (Education Trust, 2007).

It is not the first time that such proposals have been made. In fact, the impulse for national standards and assessments has been present from the beginning of the standards movement. The 1994 Improving America’s Schools Act, for example, included provisions for a new federal agency that would have certified a set of voluntary national content and performance standards against which state standards could be compared and also certified. This provision was soon repealed, however, in part because opponents perceived it as federal intrusion on state educational authority. Currently, NCLB only approves states’ standard-setting processes, not the substance of the standards themselves. In place of the federal government, many other national organizations have tried to press states to adopt more challenging standards and assessments through cajoling—public reviews and critiques of standards—or collaborative development ventures.

In many ways, these initiatives can be seen as efforts to offset the political accommodations that often accompany efforts to satisfy diverse constituencies and build consensus for education goals. Such accommodations produce content standards that are vague or extremely numerous (Massell and Kirst 1994). Furthermore, political, social and educational considerations also mitigate against the adoption of challenging performance standards for student and school accountability. Current requests for common standards are prompted by evidence that, despite public critiques and collaborative initiatives, many state performance standards are low when compared to NAEP, and state content standards are insufficiently specific, overly broad in scope, or focus on low-level skills (Finn et al 2006; AFT 2005; Wilson and Berenthal 2005).

In our interviews with policymakers, we asked questions to determine their views on participating in endeavors to create common, voluntary content and performance standards. A small number expressed cautious optimism that such an effort could be worthwhile if it resulted in high standards or if was targeted to specific problems like the Limited English proficient assessment and accountability issues mentioned earlier. In fact, four of the states here had recently or were currently participating in some standards-related initiatives. For example, Massachusetts had worked with Achieve, Inc to review its mathematics standards, and others reported reviewing other standards documents to inform their own standards development, such as Florida's use of the Fordham Foundation's A-rated standards. Texas and Massachusetts participated in the American Diploma Project to align high school standards, curriculum and assessment with college and workplace readiness.

But most of the 21 policymakers interviewed for this report were not enthusiastic about the idea of common national standards, and were not confident that developing common standards would produce good results, either politically or in terms of achieving more rigorous standards. Said one Massachusetts official: "Intellectually, it is something worth striving for. But it is complicated politically and it would require a great deal of political capital.... There is always apprehension about these discussions because every time you re-open the conversation you are faced with the possibility of more contention. We are happy to have those battles behind us." And policymakers in Massachusetts, Florida and California all thought that their own content and/or performance standards

were challenging relative to most states in at least some grades and subject areas, and feared that tying themselves to other states might lower their standards. Only one person cited NAEP comparisons, so prominent in many national dialogues, as a reason for engaging in a common standards effort. This North Dakota official thought that using the NAEP mapping study to build an index of state achievement would alleviate the local political pressure of setting cut scores “which was extremely painful before.”

However, even when policymakers were critical of the level or content of their own standards, they were equally skeptical that a common standards effort would produce challenging results. For instance, the strongest critic of Texas standards we interviewed—one who argued that standards-based reform had blanketed the state with mediocrity—also surmised that cross-state efforts to create consensus would lead back to weak standards.

Many were also concerned that common initiatives might be captured by “fads” or special interests. In fact, one California official said that the governor’s office had once vetoed funding for state participation in external collaborative efforts out of concern that the state would embrace new fads. Said one Texas policymaker, “We would not have different laboratories to see what works and what does not. I favor multiple opportunities...because what if we are wrong?” Finally, policymakers from nearly every state voiced concern was that common standards would homogenize content and remove elements that were unique to their own histories and values.

These reactions from this sample of policymakers suggest that despite the problems and challenges that states encounter in developing and maintaining unique standards, they are skeptical that voluntary, common standards would alleviate the problems and prize their own efforts, however flawed they might be.

Conclusion

The findings reported here suggest that the framework of standards-based reform continues to receive widespread support from many segments of the population, and is a stable, central part of the policy landscape in mature reform states. Policymakers and educators share the values of these reforms—especially, the goals of more rigorous

education for all students. But standards reforms encounter many difficulties in realizing their full potential, difficulties often related to the capacity of the system to provide the necessary supports for educators. Another difficulty lies in developing and adopting truly challenging standards. While many see their current standards as too low and have discussed raising them for high school students and for the accountability system as a whole, doing so incurs political, social and educational costs. And while common voluntary initiatives might help ease these problems, states in this country strongly guard their autonomy and their freedom to experiment with educational change. Furthermore, just as setting standards, assessments and accountability is by itself an insufficient remedy for improvement within states, such proposals will not address capacity and resource differences across states. Previous efforts have had varying levels of success, and comparative research on how and why different initiatives have accomplished their goals or failed may be useful.