

THE EYE OF THE BEHOLDER: SUPERINTENDENT PERCEPTIONS OF STATE SYSTEMIC REFORM

Introduction

Since their inception, one of the major conceptual bases for standards-based reforms has been the notion of alignment. In a local sense, alignment means that in order to improve student achievement and learning, school districts need to have clear standards of attainment with which their curriculum and accountability measures must align. At the state level, this notion of alignment or coherence refers to the state-developed policies that comprise its accountability system. In the most basic way, this would entail two main elements—the content standards themselves and the assessments used to evaluate school and student performance. As with the local level, this notion of alignment is paramount and the logic straightforward: the greater the degree of alignment and coherence between the state standards policies and the assessment measures, the greater the power (strength) of the accountability system, and thus, the greater the likelihood of seeing significant improvement in student outcomes. Following closely on this alignment issue is that of policy strength or impact. Again, the logic is straightforward: in essence, the bigger the stick, the greater the motivation to act and improve. Thus, it seems reasonable to believe that states that evince high-impact, high-coherence policies should also evince more activity at the local level and, ultimately, greater gains in outcomes for students.

This study attempts to investigate the first part of this proposition by examining superintendents' perceptions of state policy strength or impact in two states, one that exhibits what can be termed high impact policies (New Jersey) and the other that exhibits relatively low impact policies (Pennsylvania). Previous work has examined the intent of systemic reform policy formation in both of the states, searching for intended policy outcomes and explicating the means by which each state hoped to achieve the desired ends (Prestine, 2003). This study now moves to the next step: an examination of the perceptions and understandings of the first-line implementor of such policies—superintendents.

A brief background review of two key factors, systemic school reform and policy implementation issues, is presented here as a means to orient the findings from this study.

Systemic School Reform

Since its introduction in the early 1990s, conceptions of systemic school reform have grown and evolved. For all intent and purposes, many people now act as if systemic school reform is synonymous with standards-based reform, and the terms are used interchangeably. However, it is the original ideas of systemic school reform that are of interest here as something somewhat separate from the more recent and more narrow incarnation

as standards-based reform. Since these ideas are foundational for the study outlined below, it is appropriate to revisit them.

As first articulated by Smith and O'Day (1991), systemic school reform offered a fairly radical approach to changing and improving schools. The authors argued forcefully that it was little wonder that schools were in the sorry state that they were because the system was, in fact, in total and complete disarray. Initiative after initiative, policy after policy, program after program had come down the pike only to meet with dismal failure, especially in the nation's large urban schools. Educational policy was at best a hodgepodge with no rhyme or reason to it, and local districts and schools were proving to be largely intractable to efforts to significantly change them.¹ A growing number of critics (see, for example, Clune, 1993; Fuhrman, 1993; Timar, 1989) pointed out that many of the problems were directly attributable to an appalling lack of policy coherence both within and between levels of the educational system. Local educational agencies (LEAs), limited by position and power within the system, were simply incapable of meaningfully addressing or rectifying the situation. States, Smith and O'Day argued, long neglectful of their duties and responsibilities toward education, now had to assume a leadership position.

What they proposed was a two-pronged approach for reform. The first prong, and the one of primary interest here, was each state's responsibility to create what they called "a coherent system of instruction guidance" (Smith & O'Day, 1991, p. 20). This would bring together a fragmented system by coordinating three key functions affecting instruction—curriculum, professional development (both preservice and inservice), and assessment—and instituting mutual accountability practices in relation to them. The second prong focused on a re-examination of governance structures at all levels of the system so that, as Smith and O'Day (1990) noted, "all levels operate in support of each other and of the implementation of the reforms" (p. 20). They argued that once states overcome the fragmentation that then bedeviled the system, by providing the superstructure for overarching instructional goals, materials, professional development/training, and assessment, then the governance of the system could largely be decentralized to allow LEAs to restructure according to principles of site-based management, teacher empowerment, and professionalism. Thus, the first prong proposed a tightening of state authority or centralization on key elements of the educational system (curriculum, professional development, assessment, and accountability) while the second prong proposed a decentralization of the governance system that would allow and support LEA efforts to achieve state-set goals.

Policy Implementation Issues

The potency of states' standards-based systemic reform policies to change the educational practices of districts, schools, and classrooms likely depends on a number of key factors. Most critically, and as with any policy promulgation, there are always the intended effects of the policy, and then there are the actual effects as the policy is implemented and experi-

enced. Between these two, wide gulfs can emerge. A booming cottage industry has grown and flourished for at least the past thirty years focused on examining this gap and the reasons for a lack of fidelity in policy implementation and, concomitantly, a lack of attendant results from said policy.² Researchers documented numerous explanatory factors in an endless variety of different permutations across cases. As Berman (1981) wryly noted, “There are many ways to fail, but few to succeed” (p. 255).

A dominant explanatory theme that has emerged from this literature is that of *mutual adaptation*. In this view, the policy enters the educational system with certain intents or ends to be achieved. The focus then is on the interaction of that policy with the contexts it passes through at each stage—from state to district to school to classroom. This perspective argues that strong and prevailing contextual factors exert control over understandings and interpretations of the policy. The end result is that the policy is changed as much by the process as it changes the settings toward which it was directed, resulting in a mutual adaptation of the policy and the context to each other (Berman & McLaughlin, 1975). This explained the wide variation witnessed in policy implementation across schools and districts and the dilution of any large-scale policy effects. It also focused attention on the contextual factors that appeared to make the most difference in the implementation process—the capacity of LEAs as implementors as well as their will to put the policy into practice (McLaughlin, 1987).

As useful as this perspective has been for understanding the effects of local contexts on state policy implementation, it may overestimate the power of LEAs to modify policy and underestimate the power of the state to force compliance. As well, troubling questions about the exact meaning of the key concepts (will, capacity, mutual adaptation) persist in spite of decades of work in this area. Finally, conceptions of mutual adaptation developed around the study of policy initiatives that were, for the most part, project-oriented, aimed at specific, discrete elements of the educational system, and relatively short-term. These initiatives did not incorporate a larger view that considered the systemic forces and interconnected conditions that influence actual practice. As McLaughlin (1991) noted, such narrow project-oriented reforms “frame the problems of reform artificially and superficially and so are limited in their ability to significantly change educational practice” (p. 153). Attempting to apply a framework developed from studies of project-based, fragmented reform initiatives to examinations of systemic school reforms will likely lead to inappropriate and, worse, misleading conclusions.

New Orientations

In the last few years increasingly persuasive arguments have been made that (a) systemic state accountability reform policy is best understood as a whole rather than as discrete, individual policies, and (b) that in setting state content standards and holding districts and schools directly accountable through state-imposed assessments, state policy makers are having a deep impact on local school practice by setting the agenda and

narrowing local programmatic choices and outcomes (e.g., Firestone, Mayrowez, & Fairman, 1998; Malen, 2003). These perspectives represent a major departure from previous understandings that argued for the relative impermeability of the local site and the multitude of ways and means “street-level” bureaucrats had at their disposal to blunt the thrust of external policymakers (Weatherly & Lipsky, 1977). As Malen and Muncey (2000) noted: “State policies appear to be penetrating schools, shaping site priorities and practices, and precipitating changes despite site actors’ reluctance, resentment, and resistance. Through various combinations of symbols, sanctions, rules, regulations, and exhortations, state policies seem to be creating ‘a new set of givens’ that limit the latitude of site actors” (p. 229). If this notion of high impact state policy bundles is correct or even close to the mark, this means that an agenda for districts and schools has been established and should now be evidenced in strategies and activities of districts and schools throughout the state systems.

New Conceptions of Systemic Reform as “Policy Bundles”

The past decades have witnessed a flurry of activity by scholars examining both policy selection and implementation processes (see, for example, Clune, 1993; Elmore, Abelman, & Fuhrman, 1996; Elmore & McLaughlin, 1988; McLaughlin & Shepherd, 1995; McDonnell & Elmore, 1987). McDonnell and Elmore (1987) called attention to the fact that analyses conducted in the 1970s and 1980s tended to rely on methods that were either too narrowly focused or too empirical to capture critical contextual factors. Thus, these studies failed to account sufficiently for the complexity indicative of the policy process as a whole. Policymakers and practitioners, in general, tend to lean on different sets of considerations when, respectively, crafting policy and implementing policy. For policymakers, whom McDonnell and Elmore regard as facing issues in a “real world” political context, the tendency is to view policy-building as a highly pragmatic exercise consisting mostly of fiscal cost deliberation and political posturing to gain or preserve constituency support. Practitioners, on the other hand, tend to center their interest on implementation and accountability issues and, of course, on outcomes. As Elmore and McLaughlin (1988) aptly noted, “Elected officials and high-level administrators can reap the rewards of reform by initiating; practitioners can reap the rewards, if at all, only by implementing” (p. 59).

While the literature is replete with specifics of successful or (more likely) unsuccessful attempts at implementation for a given case or cases, little attention has focused directly on the role the state policies play in implementation. Given the existing surfeit of state systemic reform policies, let alone the breath-taking requirements of the federal “No Child Left Behind” legislation (2002), it seems imperative, in order to better understand the complexities and interactions inherent in systemic reform efforts, that the policies be thought of as “bundles” rather than individual, stand-alone policies.

In examining policy and policy types, McDonnell and Elmore

(1987) argued for a “parsimonious” yet clearly delineated analytical framework for examining and explaining the connections between policy intent and the policy mechanisms used. They suggested a typology of four generic policy types—mandates, inducements, capacity-builders, and system changers—as a means of assessing a policy’s probable impact on implementation, expected results, costs and benefits, and the likelihood of achieving its intent. Of these policy types, mandates and inducements tend to be the policy instruments of choice at the state level and, as such, they represent the “carrot-or-stick” approach to change. Mandates are the “stick.” They tend to be highly regulatory in nature, imposing rules on and demanding compliance from the policy “targets.” Inducements, on the other hand, are the “carrot.” They proffer sweet rewards (often money) to the target organization in return for specified behaviors, actions, and/or results. The other two policy types are much less common. The capacity-builders refer to policies that emphasize long-term investments in building human, material, or social capital. Some (though not all) professional development programs would serve as an example. Perhaps even more rare, system-changers take aim at altering existing patterns of delivery of services and authority relationships. One of the most prominent examples of a system-changer policy would be educational vouchers.

Elaborating on this framework, McDonnell (1994) suggested that perhaps a fifth policy typology also exists—what she terms, “hortatory” policy. This policy type, rather than drawing on authority, incentives, or capacity, relies more on persuasion. Paraphrasing McDonnell (1994), hortatory policies provide information, signal that particular goals and actions are considered high priorities by government officials, appeal to values and beliefs, and rely on positive images. While carrying no direct enforcement component themselves, they are almost always also linked to other policies that do have direct consequences for noncompliance. Thus, the defining characteristics of hortatory policy are its reliance on persuasion and the necessity of linking it with other policy instruments in order to produce the desired results.

It is with this inclusion of the fifth policy typology that the notion of a policy “bundle” is born. This notion of a “policy bundle” posits a new way to understand not only policy intent but also the more elusive implementation and impact. In addition, this concept of state systemic reform policy as a policy bundle fits well with conceptions of coherence, consistency, and complementarity—all very important features of systemic reform policy. As McDonnell and Elmore (1987) noted, “A conceptual framework focused on policy instruments not only holds the potential for moving beyond static descriptions of the implementation process, but it also embeds key variables such as local response patterns in a larger, theoretically richer context” (p. 135).

Analyzing either implementation or impact of singular policies in an era of systemic reforms is not likely to be a particularly productive or useful activity. The notion of the policy “bundle” is more appropriate. Malen (2003) would seem to concur when she notes, “Through *various combinations of symbols, sanctions, rules, regulations, and exhortations*

[my emphasis], state policies seem to be cultivating multiple direct and indirect avenues of influence on schools. . . . Perhaps states have discovered, or are in the process of discovering, *sets of policy instruments* [my emphasis] that enable them to tackle the ‘loose coupling’ tendencies of the educational policy system and ‘tighten the grip’ on public schools” (p. 209).

New Perspectives on State Policy as High-Impact

As noted above, there is little doubt that the states’ systemic reform policy bundles are aimed at significant changes in districts, schools, and classrooms. Yet, as has also been evidenced, other policies touting goals and intentions of comprehensive change have largely come and gone with little evidence of substantive, widespread, or lasting change. Malen and others (Malen, 2003; Malen & Muncey, 2000) have put forward a provocative thesis that the nature of the systemic state policy bundles marks them apart from previous policy predecessors. They argue that these new standards-based reforms may possess the potency to actually surmount local factors that have supported the status quo—the maintenance tendencies of districts, schools, and classrooms that have substantively diluted other policy attempts at change. This is largely because of the sheer force of volume, comprehensiveness, and prescriptiveness of new state-driven reforms. They propose that these new systemic or standards-based reforms significantly depart from the previous generations of state educational reform efforts that proved so ineffectual and netted so little overall gain. This, Malen (2003) argues, requires one to take what she calls a *high-impact* view of these new standards-based reforms rather than the more familiar *low-impact* view that has long dominated the policy implementation literature.

For the last thirty years, the dominant view of state educational policy has been what Malen (2003) calls low-impact. Starting at least with the RAND Change Agent studies of the 1970s and continuing well into the new century, educational policy analysts have generally bemoaned the inability of the state policy, along with a multitude of other reform initiatives/programs, to effect substantive change at the local level. The reasons proffered for this ineptitude and impotency vary. They range from explanations focusing on the native ingenuity, guile, and subterfuge of local agents slipping the bounds of a restrictive bureaucracy in order to better perform their own jobs, as exemplified by Weatherly & Lipsky’s (1977) conception of “street-level” bureaucrats, to macro organizational frameworks that look to institutional theory explanations of loosely coupled and decoupled structures and conceptions of logics of confidence and legitimacy at work to buffer weak core technologies (Meyer & Rowan, 1977, 1983; Weick, 1976). At least in educational studies of implementation efforts, one of the most frequently invoked explanations is the mutual adaptation view, previously noted and largely attributable to McLaughlin (1987, 1991; see also Berman & McLaughlin, 1975). In this view, the “contextual factors exert ultimate control” over interpretations of the policy and “the upshot for policy is that as much as it may influence the set-

tings toward which it is directed, it also is changed in process” (Knapp, 1997, p. 251). This process is characterized by a mutual adaptation of the policy and particular setting or context to each other. According to this rather pessimistic view of policy implementation, the best policy designers/makers might hope for is minimal tinkering with original policy intent; the worst would be complete subversion. Regardless of the particular explanation (and there are others as well) proffered, it seemed clear that the bottom line for all was that in any face-off between state policy and local agency, the local was likely to win out eventually.

Again, regardless of the particular spin or orientation taken, the low-impact perspective dominated the implementation literature for some time and provided a reasonably explanatory model. However, Malen (2003) suggests that things drastically changed with the advent of the era of systemic educational reforms. While state activism in the educational policy arena has been very much in evidence since the 1980s, both the tone and tenor of this activity changed with the advent of systemic or standards-based reform in the early 1990s. States grew considerably bolder in taking the initiative in educational policy making and now regularly venture into areas of schooling, i.e., the development of content standards, curriculum frameworks, testing programs, and accountability programs, long viewed as falling under the purview of local educational agencies. According to Malen (2003), we are beginning to see scattered evidence that these incursions are, in fact, having an effect and are influencing schools “in numerous direct and indirect ways that may or may not be consistent with the stated aims of the policies” (p. 205). She notes that in spite of hardly any systematic investigations of the phenomena, there is already evidence of some impact on several aspects of schooling, including: curriculum content, use of time, allocations of personnel, professional development, school improvement efforts, and normative understandings of the purposes of schooling.

Problem for Investigation

If the above thesis is correct or even close to the mark, then we stand on the cusp of an imminent sea change both in our investigations and understandings of reform policy implementation and impact, and in the governance relationship between the state and local educational agencies. What this suggests is that state-level standards-based policies may in fact be exerting powerful influences on districts and schools in entirely new and different ways—ways that our old, familiar analytic lenses may not be capable of detecting. Thus, it seems imperative that we move immediately to conduct systematic and sustained studies that will map the impact of these state policy bundles on reform implementation efforts and outcomes at the district, school, and classroom levels. More complete pictures are needed of the actual effects (both direct and indirect) of state policy in order to assess whether and which state policies have been able to ratchet up the quality of learning in schools, in what ways, and at what costs.

This study begins the exploration and mapping of this terrain. As

one way to address the overall question—what has been the impact of state systemic reform policies on implementation and outcomes at the district, school, and classroom levels?—the researcher collected and examined district superintendents' perceptions of the impact of state systemic reform policy. This study answered the following question: How do superintendents in two states, one identified as high-impact and one as low-impact, perceive and understand state systemic reform policies? While the ultimate goal of such policies is to improve student learning outcomes and achievement gains, such an assessment is difficult, if not impossible, at this early stage. It has long been acknowledged in the literature (see, for example, Elmore, 2002; Massell & Goertz, 2000; McLaughlin, 1991; Porter, Smithson, & Osthoff, 1994; Spillane, 1994) that establishing causal relationships in education is a notoriously difficult and hazardous task. Added to this, the policy bundles under scrutiny in this study are relatively new agents on the scene, and districts and schools have yet to construct their responses to these new demands. To search at this point for changes in student results would seem premature. Nevertheless, the strong systemic focus of these policies would lead one to expect any such effects to show up first in changes in district beliefs, behaviors, and practices as this level interfaces most directly with the state policy level.

The argument has been made that systemic reform efforts at the state level have changed the landscape of policy implementation and impact; that, in fact, we have moved from low-impact to high-impact policy effects. Superintendents are at once the local agents most interactive with and connected to the state level as well as critical decision-makers and shapers of local response. With this mind, this study examined two sets of issues. The first dealt specifically with an assessment of the strength/impact of state-level policy. Specifically, the following questions were addressed: How do superintendents in these two states understand and interpret the strength/impact of state systemic reform policy? Why do they see it in this way? Are some elements of the policy more visible and influential than others? How do they assess the strength/impact of the standards-based reform policy bundle (high-impact/low-impact)?

The second prong of this investigation identified and examined specific areas of impact and outcomes. Specifically, the question addressed was: Do specific areas of schooling appear to be more directly impacted by the systemic reform policies than others? Malen (2003) identified several areas that she hypothesized might be the most susceptible to high impact policy bundles. These areas included: (a) content of the curriculum, (b) use of time, (c) allocation of personnel, (d) decisions about professional development, (e) intensifying work, (f) decisions regarding school improvement plans, and (g) the growing importance of testing as “the” means of assessing performance. Each of these specific areas were examined.

Conceptual Framework

To address these questions, I borrowed and adapted a conceptual framework first developed by Porter, Flooden, Freeman, Schmidt, and

Schwille (1986) at Michigan State University and later modified and used by Tyree (1993) to analyze the “strength” of intended state policies to affect curriculum content decisions at the local level. Tyree (1993) examined the inherent or potential strength of intended state policy outcomes from the perspectives of the state education agency (SEA). I modified this slightly and looked to assess the perceived strength of the actual policy bundles as experienced by the local education agency (LEA).

The framework developed by Tyree (1993) identifies four main characteristics that, taken together, offer a means of assessing the relative strength of a state policy—consistency, prescriptiveness, authority, and power. A brief explanation of each follows. *Consistency* refers to the degree of alignment or coherence between and among individual policies within the systemic or standards-based reform policy bundle. Policies should reinforce as well as compliment each other. An example of reinforcement would be regulatory policies (required state testing programs of content standards) that back up and give impetus to implement hortatory policies (curriculum frameworks/content standards). A lack of alignment is viewed as weakening the strength of the policy bundle as it sends mixed and confusing signals to the local level as to what is important and valued. An example of this is when content standards and curriculum frameworks emphasize higher order thinking skills but assessment instruments test only basic recall.

Prescriptiveness refers to the policy’s depth of detail, its specificity. This pertains to the extent that the policy gives the LEA a clear understanding and outline of exactly what is expected. The implication is that the greater the clarity of required response by the LEA, the greater likelihood that response will be forthcoming. *Authority* refers to identified and recognized legitimate sources and/or legal bases for rule/policy promulgation and expected acquiescence. The basic premise is that the more a given state policy appeals to several bases of authority, the greater the chance that LEAs will accept that policy—thus, the greater the policy strength. Appeals to authority may include the endorsements of professional organizations (teachers’ unions, administrator associations, school board associations), the use of content teachers from around the state to assist in developing state assessments, and governor/state legislature endorsement.

The last of the four characteristics is *power*. Power is distinct from authority as there is a move from the hortatory to regulatory. Power refers to the granting of rewards and incentives for those who perform well and punishments or sanctions for noncompliance or unsatisfactory performance. One of the clearest examples of this aspect is the high-stakes testing aspect of state-mandated assessments. Whether tied to grade promotion, graduation, or simply publication in the local newspaper, these tests and the content they contain are much more likely to catch the attention and time of teachers and students if important consequences are tied to them.

This seemed like an especially propitious model to adapt for use in this study for at least two reasons. First, this framework rests on the key assumption that the “stronger” the state policy, the greater the likelihood

that local agents will not only implement it but implement it with greater fidelity. Thus, the likelihood that the policy bundle will have an impact on outcomes is increased. While there are certainly other factors that influence policy implementation, not the least of which is the capacity of the LEA, this major proposition fits well with the key assumptions undergirding systemic reforms. Second, this framework has been validated for similar analytic purposes in previous studies. Tyree (1993) used this particular framework as an analytic tool for his research examining intended (as opposed to implemented) curriculum policy across several states. Clune (2001) also used it in his analysis of secondary case study data gathered from the Statewide Systemic Initiatives (SSIs) program funded by the National Science Foundation.

Methodology

State Selection

A comparative study of two states (New Jersey and Pennsylvania) was completed in 2004 focusing on superintendents' perceptions of the degree of strength/impact of state policy bundles. At least two independent databases rate these two states as quite different in their approaches to systemic state accountability policies. These two include: (a) the Achieve group's Measuring Up project (2002) and (b) a scale developed by Carnoy and Loeb (2004) that ranks states on the stringency of their accountability measures. Both are briefly detailed below.

Achieve group's Measuring Up project. Overall, the goal of Achieve (2002) has been to help individual states to assess their standards and alignment by using a common set of criteria. They have done this both by comparing state standards to sets of exemplar standards—those determined by content area groups like the National Council of Teachers of English (NCTE, 1996) and the National Council of Teachers of Mathematics (NCTM, 2000), and those set by other countries like Japan.

In reviewing state content standards, Achieve looks for several factors they consider crucial. These include the rigor of the standards, the degree of comprehensiveness and consistency, and their clarity. Using a benchmarking process, comparisons are made between the state's standards and the exemplars. After this, the state's assessment is examined to determine whether or not it measures what the standards require. In other words, they seek to answer the question—do the content standards align with the tests being used to assess them?

The alignment issue is critical as these two important prongs of the state accountability policy bundle (standards and assessments) must complement and reinforce each other if either is to be effective. This alignment issue speaks directly to Tyree's (1993) conceptions of consistency and prescriptiveness. Since the degree of alignment is a result of the relationship between the content standards and the assessment instrument used, it could reasonably be improved by changing either (Webb, 1997).

While over 20 states, including both Pennsylvania and New Jersey, have undergone an Achieve review, not all this information is directly available. A full report of New Jersey's review is available, but the Pennsylvania review is not, making direct comparisons difficult at best. Nonetheless, it is important to note that in reviewing New Jersey's content standards, Achieve (2000) noted some problems. However, most of the problems noted in the Achieve report have since been rectified. As well, New Jersey's assessments (with the exception of the high school exit exam) have consistently been rated as exemplar. Pennsylvania, on the other hand, has consistently trailed the field in the development of both content standards and state assessments. The very fact that Pennsylvania declined to make its Achieve report public may speak volumes.

State accountability scale. Carnoy and Loeb (2004) developed a means of assessing the strength of each state's accountability policies based on the database developed by the Consortium for Policy Research in Education (CPRE).³ Using this database, they constructed a scale of accountability levels ranging from 0 to 5, with zero representing states that had no statewide standards nor state assessments. States at the high end with a score of 5, on the other hand, had students tested in several different grades, strong sanctions and rewards for schools based on student test outcomes, and a high school exit exam required for graduation. The remaining scale scores represented shades of variation between these.

At the time that this scale was developed, New Jersey, along with Texas and North Carolina, scored a 5 on this index. Pennsylvania scored a 1, indicating that it required testing to state standards only in lower grades but had no school sanctions or rewards (weak repercussions), and no required high school exit exam.

Superintendent Survey

The surveys were mailed in 2004 to approximately 240 school districts across the two states (approximately 20% of the districts in Pennsylvania [$n = 120$] and New Jersey [$n = 120$]). The districts were randomly selected. The survey was addressed to superintendents or chief executive officers. Overall, the survey asked the same questions of superintendents from both states although minor wording changes were necessary to make some of the survey items state-specific. The survey was constructed around two key intents. The first sections of the survey were designed to gauge policy strength or impact using Tyree's (1993) framework for assessing curricular policy. Survey items were specifically designed to assess how superintendents perceived the strength of the state policy bundles on their district as measured by each of the concepts of consistency, prescriptiveness, authority, and power. All items were of a Likert-type using a score of 5 to indicate "strong agreement" and 1 to indicate "strong disagreement." The second section of the survey moved from a focus on the state policy bundles themselves to their implementation in and impact on the LEA. Malen (2003) identified six specific areas that she

theorized may be impacted the most by state systemic policy. Respondents were asked to rank order these areas from 1 (most impacted) to 6 (least impacted). Open-ended questions were provided in all sections of the survey for further comments and clarifications.

Return rates for both Pennsylvania and New Jersey were low, with 47 useable surveys returned in Pennsylvania for 39% return rate and 35 useable surveys returned from New Jersey for a 22% return rate. Seven surveys from Pennsylvania were returned only partially completed and 5 from New Jersey. These surveys were not used in the analysis. Again, it is likely that the timing of the survey (March/April, 2004) was at least in part to blame for the low response. Because of the low response rate, only summary data statistics are reported here.

Findings

New Jersey

As might be anticipated, superintendents from New Jersey perceived their state policy as having a high impact. Table 1 offers data from the survey concerning the assessment of the strength or impact of the policy bundle. Superintendents considered the prescriptive nature of the policies as especially strong and high impact. Not only did the superintendents rank this as the highest of the four categories, it also had the lowest standard deviation of any of the other subcategories. As one of the superintendents commented, "We live in a take no prisoners environment. The state has made the consequences for less than satisfactory results abundantly clear." Another noted, "Thankfully, I retire next year. I don't know how anyone is going to be able to navigate this kind of pressure-cooker."

Table 1

Mean Scores and Standard Deviations From New Jersey for Policy Impact Subsections

Consistency	Prescriptiveness	Authority	Power	Total impact
4.09 (1.0015)	4.76 (0.428)	3.60 (1.1764)	3.86 (0.729)	4.15

N = 35.

The responses of the superintendents when asked to rank order the changes made in response to state policy mandates were revealing. The unanimity found in the previous section continued to the ranking of change areas. Table 2 presents the data from the rank ordering of the six items. Again, a ranking of 7 indicated an area that has been most impacted by state policy and a ranking of 1 represents an area least affected. An average score for each item is given.

Table 2*Mean Scores From New Jersey for Rank Ordering of Change Areas*

Content of the curriculum	5.58
Use of time	4.20
Allocation of personnel	3.73
Professional development	3.01
Intensification of work	2.77
School improvement plans	3.29
Outcome assessments	2.21

N = 35.

Clearly, the respondents viewed the content of the curriculum as being the area most impacted by state policy. With the relatively tight alignment that exists between New Jersey's state standards and the state assessment, it seems reasonable that this area would be most directly affected. Few respondents took the opportunity to supply additional written comments for this section. However, one respondent noted in response to the "intensification of work" item that "I don't think we are working harder; we are working smarter." Another respondent noted in referring to the "outcome assessments" item, "There is *only one* assessment that matters" (respondent's emphasis). It seems clear that all these areas have been affected to some degree or another. Whether there are meaningful differences in the rank ordering of the items, however, remains questionable. There was a fair amount of unanimity in the ranking of items. However, given the low return rate, any further interpretation must be suspect.

Pennsylvania

The pattern of responses from Pennsylvania superintendents echoed in many ways the responses of New Jersey superintendents but with a few important differences (see Table 3). With 47 useable surveys returned, Pennsylvania superintendents also rated the impact of state policy bundles as fairly high, especially in the area of prescriptiveness. However, the standard deviations for all categories were fairly large, indicating a general lack of agreement or consensus. It also appears that the terms used (consistency, prescriptiveness, authority, and power) may have different meanings and associations for Pennsylvania superintendents. In the written responses for this section the Pennsylvania respondents consistently referred to the No Child Left Behind (NCLB) requirements rather than referring to state policy mandates even if they did not identify these as such. For example, one respondent noted, "AYP [Average Yearly Progress]—that's what impacts us. You should have this as a category." Another noted, "All of these apply to NCLB. That's the whole point." There appears to be no real differentiation between state policy actions and the federally mandated NCLB legislation for these administrators. NCLB was in fact what respondents consistently referred to and it did not

appear to be distinguishable from state policy mandates in any meaningful way for the respondents.

Table 3

Mean Scores and Standard Deviations From Pennsylvania for Policy Impact Subsections

Consistency	Prescriptiveness	Authority	Power	Total impact
4.08 (1.635)	4.22 (1.203)	2.63 (1.917)	3.81 (1.395)	3.80

N = 47.

In responding to the rank ordering of the areas of impact of the state policy initiatives, the Pennsylvania superintendents again revealed a lack of unanimity. As Table 4 reveals, other than the top choice for areas impacted by state policy (that of content of the curriculum), little if any agreement was exhibited. This is clearly evidenced by the relatively similar means that were computed for each of the change areas.⁴

Table 4

Mean Scores From Pennsylvania for Rank Ordering of Change Areas

Content of the curriculum	5.17
Use of time	3.64
Allocation of personnel	3.16
Professional development	2.35
Intensification of work	2.42
School improvement plans	2.47
Outcome assessments	2.86

N = 47.

Written comments again were sparse for this section. Some of these again referenced NCLB while others focused on rural school issues. One respondent noted, “No one really understands or cares about how any of this hits rural districts. There is a real lack of understanding in Harrisburg (the state capital). Where do they think these kids are going to go if the nearest district is over an hour away on the other side of the mountain ridge?”

A Comparison

Based on the results from this limited study, the assertion that high impact policy will be perceived differently than low-impact policy appears to have some justification. As can be seen in Table 5, a direct comparison of the two states reveals that superintendents in New Jersey, a high-impact policy state, consistently rated each of the policy impact areas higher than did their counterparts in Pennsylvania, a low-impact

policy state. The determination of whether these are in fact significant differences and a closer examination of variance differences remain for future research efforts. Nonetheless, Tyree's (1993) framework appears to be sensitive enough to pick up differences in perceptions of the strength or impact of state accountability policy by the local level.

Table 5

Comparison of Mean Scores and Standard Deviations for Policy Impact Subsections

	Consistency	Prescriptiveness	Authority	Power	Total impact
Pennsylvania <i>n</i> = 47	4.08 (1.635)	4.22 (1.203)	2.63 (1.917)	3.81 (1.395)	3.80
New Jersey <i>n</i> = 35	4.09 (1.0015)	4.76 (0.428)	3.60 (1.1764)	3.86 (0.729)	4.15

In examining a comparison of the two states' rankings of change areas, there was complete agreement on the top three areas changed because of the state policies (see Table 6). These included the content of the curriculum, the use of time, and the allocation of personnel. Consistent with being from a high-impact state, New Jersey superintendents, however, ranked changes in these areas consistently higher than did Pennsylvania superintendents. These higher ratings extended across all the areas mentioned except for that of outcomes assessments. This might be partly explained by the fact that New Jersey has a much longer history of outcome assessments whereas in Pennsylvania they are a relatively new phenomena and one that they are, therefore, acutely aware of. The other item drawing an unanimity of response from superintendents in both states was intensification of work, ranked sixth from both states.

Three items drew the largest differences in mean scores from the two states. These items included school improvement plans, professional development, and outcome assessments. Possible differences in response for outcome assessments were already noted. The area of school improvement plans showed the greatest variation in ranking. New Jersey schools have for several years been required to submit these plans regardless of their performance on the state assessment. However, in Pennsylvania, school improvement plans are again a relatively new and limited phenomena. School improvement plans are not required of schools or districts unless they fail to meet AYP expectations for two consecutive years. This may at least in part explain the much lower ranking given to this area by Pennsylvania superintendents. In a similar fashion, professional development assumes a much different look in the two states. New Jersey, with its high-impact and highly aligned policies, ties professional development and its funding directly into its school accountability policies. Pennsylvania, on the other hand, is among the states with the lowest percentage of school funding coming from the state with approximately 30% funding

from the state. This puts the majority of professional development funding directly on the shoulders of the individual districts. In a state with a majority of districts being both rural and poor, professional development must often be sacrificed to meet more immediate and pressing funding needs.

Table 6

Mean Scores From Rank Ordering of Change Areas

	New Jersey (Ranking)	Pennsylvania (Ranking)	Difference in means	Average (Ranking)
Content of the curriculum	5.58 (1)	5.17 (1)	.41	5.37 (1)
Use of time	4.20 (2)	3.64 (2)	.56	3.92 (2)
Allocation of personnel	3.73 (3)	3.16 (3)	.57	3.45 (3)
Professional development	3.01 (5)	2.35 (7)	.66	2.68 (5)
Intensification of work	2.77 (6)	2.42 (6)	.35	2.59 (6)
School improvement plans	3.29 (4)	2.47 (5)	.82	2.88 (4)
Outcome assessments	2.21 (7)	2.86 (4)	-.65	2.54 (7)

Conclusions and Implications for Further Study

Several important conclusions can be drawn from these findings regarding the effects of high-impact and low-impact state policy bundles. These conclusions center on: how superintendents understand and perceive state accountability policy, the impact these understandings and perceptions have on actions taken and changes made, and the susceptibility or resiliency of different aspects of schooling to high-impact policies. Implications focus on a discussion of the viability and efficacy of high-impact versus low-impact policy bundles in promoting significant educational change.

Superintendent Perceptions of State Accountability Policy Bundles

The findings from this study indicate at least a modest association between superintendents’ assessments of state accountability policy and the strength of the policy bundle. In New Jersey, a high-impact policy state, superintendents rated the overall strength of the state accountability policy as high at a mean score of 4.15 out 5.0 while superintendents in Pennsylvania, a low-impact state, rated the overall strength of their state’s accountability policy at a more modest 3.80.

More significantly, New Jersey superintendents rated all subsections of the policy bundle impact as high, confirming the idea that alignment of policies increases policy strength. With the increase in policy strength there appears to come a concomitant increase in policy awareness and attentiveness to such on the part of superintendents. In Pennsylvania, where state accountability policy development has been at best hit-or-miss, there appears to be a general lack of awareness of state accountability policy. In the open response items, superintendents from Pennsylvania consistently referred to the NCLB policy, not state policy. In point of fact, other than initial work on content standards for language arts and mathematics, Pennsylvania accountability policy did not exist for all intents and purposes before the requirements of NCLB came into effect. Thus, it is not surprising that Pennsylvania superintendents refer exclusively to NCLB and not to state policy. New Jersey superintendents, on the other hand, were very much aware of the New Jersey policy as the state has been working aggressively on this policy for almost a decade. Interestingly, none of the New Jersey superintendents referred to NCLB.

Across both states, superintendents tended to view the accountability policies as a whole, although New Jersey superintendents perceived a tighter linkage than did Pennsylvania superintendents. Both groups especially noted a tight linkage between the state content standards and the assessment. Whether in fact this tight linkage actually exists may be questionable. The most recent review of accountability measures in New Jersey, already noted as a high-impact state, by Achieve (2002) indicated that persistent problems with alignment and rigor still exist. At least in some content areas, tests still do not measure what the standards expect regarding content or performance, and rigor continues to be problematic. Considerably less is known about the extent of alignment of the Pennsylvania standards and assessments.

Finally, superintendents in both states, of course, had only their state-specific experience from which to draw on to assess their perceptions of state policy strength. While external evidence from both Achieve (2000, 2002) and Carnoy and Loeb (2004) strongly suggest that significant differences in policy orientation do exist between the two states, superintendents had only their experiences within a given state to draw upon. In examining standard deviations, it seemed clear that New Jersey superintendents showed a remarkable unanimity of agreement while Pennsylvania superintendents had more diversity of response. This may be an artifact of the differences between high- versus low-impact state accountability policies. High-impact policy bundles may have a clarity that is communicated to intended audiences that low-impact policies may not.

Impact of Perceptions on Actions and Behaviors

By the very nature, low-impact policies lack the aggressive, mandatory nature of high-impact policies that tend to be tightly coupled around reinforcing systems of clearly defined rewards and sanctions. The “wobble” room afforded by such policies is slight. Critics of such central-

ized state educational policymaking argue that districts and schools experience a loss of control over the substantive areas of schooling. This concern assumes that state accountability policies exercise a level of control that can constrain districts and, concomitantly, superintendents from following their own beliefs and understandings about schooling and what is best for their students.

This kind of constraint would seem to be most likely in the case of high-impact states. It seems inevitable that there would be instances wherein superintendents would have made other choices or focused on other areas but for state policy mandates. However, even in New Jersey, a high-impact state with strong sanctions and rewards, superintendents reported that they maintained relatively high levels of control and discretion. Several of the New Jersey superintendents' comments indicated that they felt neither hamstrung nor disenfranchised by the state mandates. Several of these superintendents noted that they were already predisposed to move in the directions of the state policy. The policies themselves, based on principles of equity, accountability, and high standards for all students, had given the superintendents a leverage they did not previously have to enact changes and policies they endorsed within their districts for instructional improvement. While there was some individual variation in superintendent beliefs and understandings of these policies, on the balance the changes required were not viewed as usurping prerogatives. As one superintendent noted, "There's nothing unreasonable here. We should be doing this anyway." Another added, "I see these [the New Jersey policies] as useful tools for improving student learning."

In Pennsylvania, the low-impact state, the changes required were less clear as the state accountability policy bundle itself was less well-defined. With no substantive consequences hanging in the balance (other than meeting AYP), state policy was seen (accurately) as merely the extension of the federally mandated NCLB. The changes instituted and the actions taken devolved to the district level discretion and related to AYP demands. For all intents and purposes, there is no state policy bundle in Pennsylvania—only the efforts of the state to meet NCLB requirements. This gives districts a good deal of freedom to make changes they desire but little direction. In Pennsylvania, policy still exemplifies the loose coupling thesis (Weick, 1976), at least as much as is possible given the NCLB requirements. While this study does not have direct observation or interview data, the superintendents responding to the survey appeared to feel that they had a fair amount of discretion as to what changes should be made. Several superintendents included comments that greater direction from the state would not necessarily be a bad thing, especially if extra funding were provided to the districts. One superintendent commented, "[It is] hard to decipher what PDE [Pennsylvania Department of Education] wants. Tell us what to do and we'll do it." Another noted, "We decide, but we have to pay for it." Finally, a third superintendent commented, "[There is] little assistance from the state—they don't know how to do this either."

Susceptibility and Resiliency

It seems likely that some areas of schooling are more vulnerable to high-impact state accountability policies than are others. Data from this study indicated some difference with regard to perceptions of the areas impacted by policy. However, a caveat is likely needed here. While various areas may be more or less susceptible to accountability reforms, the degree and quality of the change required present an entirely different consideration. As Archbald and Porter (1994) noted, interpretations of policy impact have a Janus-like quality to them. Changes can be large-grained or they can be small-grained or they can be both at the same time. For example, the content of the curriculum, the number one change area affected by state accountability policy across both states, can be accomplished by crosswalking existing district curriculum with state standards. This is a large-grained change and brings the district's curriculum into technical alignment with the state's content standards. However, unless a small-grained change is also accomplished at the individual school/teacher/classroom level, it is unlikely that the large-grained change will have the desired effect. The most potent change is when there is evidence of and synchrony between both types of change.

The vulnerability of some areas of schooling to state accountability policies also opens the question of whether or not these are the right areas for change. There is growing concern that the changes already made and the improvements evidenced in achievement in high-impact states may have little to do with "real" student learning. As Carnoy and Loeb (2004) found, states with strong accountability did indeed see greater gains in test scores. However, as their data show, strong accountability had little effect in raising the retention rate of students in the first year of high school or in increasing the numbers of students who progressed through high school to graduation. These kinds of results again point to the specter of teaching to the test and improving students' abilities to be good test-takers but not significantly improving their educational attainment.

With this last point in mind, it is also important to remember that at best even high-impact state policy can only influence certain district, school, and classroom variables. Were these the only variables to influence student achievement and learning, the educational crisis would be over and the problem solved. However, as any educator knows, there are a multitude of external variables too numerous to even begin to list here, that also influence students' educational attainment. This is not an excuse that we do not need to "get it right" as far as those factors under our control, or at least partially under our control, are concerned but a recognition that ensuring educational attainment for all students is an extremely complex issue.

Implications for Further Study

The interpretation of summary data results cannot be pushed beyond their obvious limitations. These limitations refer specifically to the low-response rate for the survey and, consequently, the use of descrip-

tive statistics. This study examined state accountability policy bundles to assess differences between superintendents' perceptions of either high-impact or low-impact state policies. This was accomplished for the respondent sample, approximately 30% of the districts surveyed. However, it does seem evident that some implications for further study do emerge from this effort.

First, and most significantly, even this very limited study provides preliminary support for using the notion of "policy bundles" and the conceptualization of high-impact versus low-impact state policy as a means of investigating and understanding policy implementation, impact, and change. Necessary next steps should include: (a) conceptual clarification of the differences between high-impact and low-impact state policy bundles, and (b) further investigation, refinement, and elaboration of the "direct and indirect" means (Malen, 2003, p. 209) by which such policy bundles impact district, school, and classroom changes as well as those areas most vulnerable to such policy initiatives.

At least within this study, it appears clear that high-impact policies, whatever their limitations, do bring clarity, focus, and even energy to school improvement efforts and open areas to change that may have previously been closed or at least difficult to confront. Whether or not all this is the "right" focus and direction for all districts or any given district is another question for further study. As others (Malen, 2003; Prestine, 2003) have suggested, it simply may be too early in the life of the high-impact state accountability reforms to assess their efficacy in any meaningful way.

In spite of what looks like a positive relationship between superintendent perceptions and policy strength, other studies need to be done of other states with similarly weak or strong accountability systems. The two state analysis completed here produced some provocative findings that need further exploration, analysis, and confirmation/disconfirmation in other states identified as high-impact and low-impact. In addition to this, perceptions of teachers, principals, students, parents, and other community members would permit a more complete picture of state accountability policies. While superintendents are clearly leading meaning-managers for policy interpretation in their district, gaining the perspectives of other stakeholder groups would be most beneficial.

Finally, student achievement gains must be the ultimate test of policy strength/impact. This clearly needs to be on the agenda in any effort to examine differences between high-impact and low-impact states. The long-term effects of strong- versus weak-impact state accountability policy bundles on student achievement are still largely unknown. This is likely due to several factors. First, high-impact policy bundles have been around a relatively short period of time. Any assessment of efficacy is difficult when longitudinal data are not available. Second, state policy bundles themselves have been in a constant state of change and evolution. However, there do appear to be relatively clear state stances with regard to the inclination to go the high-impact route versus the low-impact route. Further research is needed to clarify this situation.

Notes

¹Smith and O'Day (1991), showing a good deal of discernment and foresight, correctly deduced that the restructuring movement, then the current favorite among educational reformers, would be unlikely to show the benefits so many hoped for.

²Some of the classics in this genre include Bardach (1977), Gross, Giacquinta, and Bernstein (1971), and Pressman and Wildavsky (1984).

³The CPRE database is available at http://www.cpre.org/Publications/Publications_Accountability.htm

⁴It should be noted that there may have been an over-representation of rural schools in the returned surveys. While over 60% of Pennsylvania school districts are rural and clearly represent a significant group, it appears that understandings of what constitutes professional development and issues of allocation of personnel, at least, have quite different meanings for small, rural districts that tend to be geographically isolated.

References

- Achieve. (2000). *Measuring up: A standards and assessment benchmarking report for New Jersey*. Washington, DC: Author.
- Achieve. (2002). *Measuring up: A report on educational standards and assessment in science for New Jersey*. Washington, DC: Author.
- Archbald, D. A., & Porter, A. C. (1994). Curriculum control and teachers' perceptions of autonomy and satisfaction. *American Educational Research Journal*, 16(1) 21–39.
- Bardach, E. (1977). *The implementation game: What happens after a bill becomes a law*. Cambridge, MA: MIT Press.
- Berman, P. (1981). Educational change: An implementation paradigm. In R. Lehming & M. Kane (Eds.), *Improving Schools: Using what we know* (pp. 253–310). Beverly Hills, CA: Sage.
- Berman, P., & McLaughlin M. W. (1975). *Federal programs supporting educational change, vol. VIII: Implementing and sustaining innovations*. Santa Monica, CA: The RAND Corporation.
- Carnoy, M., & Loeb, S. (2004). Does external accountability affect student outcomes? A cross-state analysis. In S. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 189–219). New York: Teachers College Press.
- Clune, W. H. (2001). Toward a theory of standards-based reform: The case of nine NSF statewide systemic initiatives. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 13–38). Chicago: National Society for the Study of Education.
- Clune, W. H. (1993). Systemic educational policy: A conceptual framework. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 124–140). San Francisco: Jossey-Bass.

- Consortium for Policy Research in Education (CPRE). (2000). State accountability systems. Retrieved from http://www.cpre.org/Publications/Publications_Accountability.htm
- Elmore, R. (2002). *Bridging the gap between standards and achievement*. Retrieved from the Albert Shanker Institute web site: <http://www.shankerinstitute.org/education.html>
- Elmore, R. F., Abelman, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65–98). Washington, DC: Brookings Institution.
- Elmore, R. F., & McLaughlin, M. W. (1988). *Steady work: Policy, practice, and the reform of American education*. Santa Monica, CA: The RAND Corporation.
- Firestone, W. A., Mayrowez, D., & Fairman, J. (1998). Performance based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95–113.
- Fuhrman, S. H. (Ed.). (1993). *Designing coherent education policy*. San Francisco: Jossey-Bass.
- Gross, N., Giacquinta, J. B., & Bernstein, M. (1971). *Implementing organizational innovations: A sociological analysis of planned educational change*. New York: Basic Books.
- Knapp, M. S. (1997). Between systemic reform and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. *Review of Educational Research*, 67(2), 227–266.
- Malen, B. (2003). Tightening the grip? The impact of state activism on local school systems. *Education Policy*, 17(2), 195–216.
- Malen, B., & Muncey, D. (2000). Creating a new set of given: The impact of state activism on school autonomy. In N. D. Theobald & B. Malen (Eds.), *Balancing local control and state responsibility for K–12 education* (pp. 199–244). Larchmont, NY: Eye on Education.
- Massell, D., & Goertz, M. (2000, April). *Local strategies for building capacity: The district role in supporting instructional reform*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- McDonnell, L. M. (1994). *Policymakers' views of student assessment*. Santa Monica, CA: The RAND Corporation.
- McDonnell, L. M., & Elmore, R. F. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis*, 9(2), 133–152.
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9(2), 171–178.
- McLaughlin, M. W. (1991). The RAND change agent study: Ten years later. *Educational Researcher*, 19(9), 11–26.

- McLaughlin, M. W., & Shepherd, L. (1995). *Improving education through standards-based reform*. (Report from the panel on standards-based reform). Stanford, CA: The National Academy of Education.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
- Meyer, J., & Rowan, B. (1983). The structure of educational organizations. In J. V. Baldridge & T. Deal (Eds.), *The dynamics of organizational change in education* (pp. 60–87). Berkeley, CA: McCutchan.
- National Council of Teachers of English (NCTE). (1996). *Standards for the English language arts*. Urbana, IL: Author.
- National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- Porter, A. C., Floden, R. E., Freeman, D. J., Schmidt, W. H., & Schwille, J. R. (1986). *Content determinants* (Research series no. 179). East Lansing: Michigan State University, Institute for Research on Teaching.
- Porter, A. C., Smithson, J., & Osthoff, E. (1994). Standard setting as a strategy for upgrading high school mathematics and science. In R. F. Elmore & S. H. Fuhrman (Eds.), *The governance of curriculum: 1994 yearbook of the Association for Supervision and Curriculum Development* (pp. 138–166). Alexandria, VA: ASCD.
- Pressman, J. L., & Wildavsky, A. (1984). *Implementation: How great expectations in Washington are dashed in Oakland*. Berkeley: University of California Press.
- Prestine, N. A. (2003, September). *Does state policy matter in standards-based reform?* Philadelphia, PA: Temple University, Laboratory for Student Success.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233–267). New York: Falmer.
- Spillane, J. P. (1994). How districts mediate between state policy and teachers' practice. In R. F. Elmore & S. H. Fuhrman (Eds.), *The governance of curriculum: 1994 yearbook of the Association for Supervision and Curriculum Development* (pp. 167–185). Alexandria, VA: ASCD.
- Timar, T. (1989). The politics of school restructuring. *Phi Delta Kappan*, 70, 265–275.
- Tyree, A. K., Jr. (1993). Examining the evidence: Have states reduced local control of curriculum? *Educational Evaluation and Policy Analysis*, 15(1), 34–50.
- U.S. Department of Education. (2002). No Child Left Behind. Retrieved from <http://www.ed.gov/nclb/landing.jhtml?src=pb>
- Weatherly, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special education reform. *Harvard Education Review*, 47, 171–197.

Prestine

- Webb, N. L. (1997). Determining alignment of expectations and assessments in mathematics and science education (NISE Brief, 1(2), 1-8). Madison: University of Wisconsin-Madison, National Center for Improving Science Education.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21(1), 1-19.

Nona A. Prestine is a Professor in the Department of Education Policy Studies at Penn State University, University Park, PA.