

***Symposium on Education Finance
and Organizational Structure
in New York State Schools:
A Synthesis***

DAVID H. MONK

AND

JAMES WYCKOFF



Education Finance and Organizational Structure in New York State Schools: A Synthesis

David H. Monk
Pennsylvania State University

James Wyckoff
University at Albany - SUNY

I. Introduction and Background

The Education Finance Research Consortium, with support from the Wallace Foundation, convened a research symposium on Education Finance and Organizational Structure in New York State Schools. The symposium was conceived against the backdrop of three important developments:

- New York's high stakes high school exit exams accompanied by learning standards that are tied to curricular changes and a revised accountability system.
- The flow of significant new dollars into the K-12 public schooling system despite New York's serious fiscal constraints.
- The New York State Court of Appeals far-reaching decision in the Campaign For Fiscal Equity (CFE) v. New York State court case. The Court gave the State until July 30, 2004 to enact finance and management reforms to ensure that every school in New York City has resources needed to provide the opportunity for an adequate education, and to measure whether reforms actually provide this opportunity.

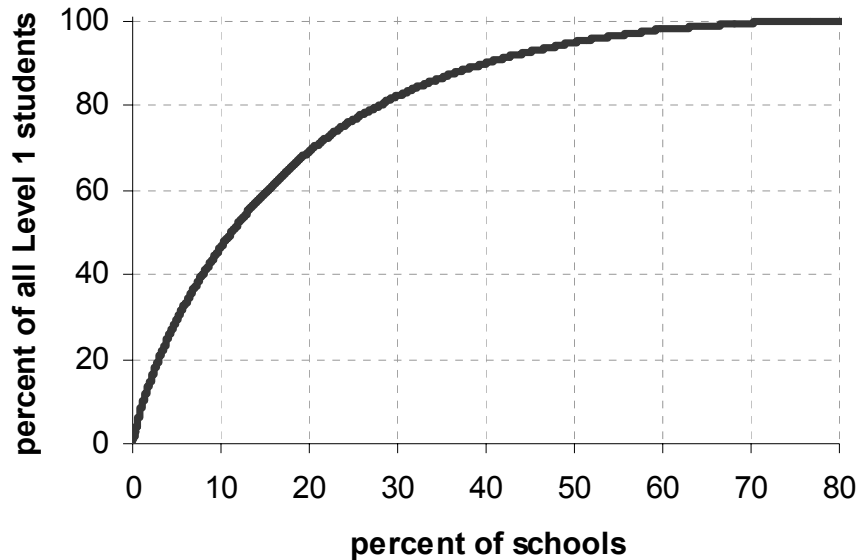
These developments prompt important questions about how New York State should develop school finance, organizational structure, incentive and accountability systems to most effectively meet the challenges of an outcomes-based approach to education and discharge its responsibilities to provide an adequate education to students throughout the State. The symposium was designed to assess evidence and develop ideas on these topics.

In particular, the symposium addressed four specific interrelated components of providing New York's school children with an opportunity for an adequate education: education finance; the teaching workforce; organizational structure and incentives; and data management. Researchers participating in the symposium were selected for their expertise in these areas and were asked to prepare policy briefs in one of these areas. Public presentation of the symposium policy briefs occurred on March 5, 2004 in Albany, New York. The policy briefs were subsequently revised in light of the discussion that took place during the public forum and appear in this volume. Here we provide a synthesis of the policy briefs and focus on the implications for policy makers.

Before turning to overviews of the individual briefs, we believe it is important to emphasize the importance of focusing attention on the questions and policies crucial to

improved outcomes for New York’s lowest performing students. Research shows that the lowest performing students are concentrated in relatively few schools and located in a handful of districts¹. In fact, 70 percent of the students who scored at Level 1 on the 2002 fourth grade math exam were concentrated in just 20 percent of the schools (see Figure 1). Seventy-five percent of these 480 lowest-performing elementary schools are in New York City; 14 percent are in the Big Four cities (Buffalo, Rochester, Syracuse and Yonkers). The incidence of very poor student academic performance in New York State is overwhelmingly an urban problem that is disproportionately concentrated in New York City. Another salient point is that even though most low performing schools are in urban districts, fewer than half of all urban schools are in this lowest performing group.

Figure 1
Concentration of Students Scoring at Level 1 on 4th Grade Math Exam
(Schools ordering based on number of Level 1 students.)



II. Overview of the Policy Briefs: Lessons from Research

Education Finance Policy Briefs

Three policy briefs examine aspects of school finance. Tom Downes compares the conceptual frameworks and empirical application of various approaches commonly employed to measure adequacy. He explores the distributional consequences of alternative methods of costing an adequate education in New York State. Jon Sonstelie examines various alternatives of raising revenues to support increased school expenditures. Leanna Stiefel, Ross Rubenstein and Amy Ellen Schwartz compare school level expenditures across schools in New York City, Cleveland and Columbus, Ohio.

¹ As reported in the Boyd, Lankford and Loeb policy brief.

Downes

In “What is Adequate? Operationalizing the Concept of Adequacy for New York,” Downes notes that adequacy is defined by policy makers and so must address either minimum resource or minimum performance requirements. If adequacy is defined only in terms of resources, then adjustments must be made for the differential cost of purchasing equal quality inputs across districts. If on the other hand adequacy is defined in terms of student performance, then it must also include differential ability of students to perform. Downes identifies the major approaches to measuring adequacy as: a) geographic cost approach; b) empirical identification approach (employed in the NYS Regents proposal); c) whole school design approach; d) professional judgment/resource cost model approach (PJRCM); and the cost function approach (CF). The empirical identification approach is employed in the New York State Board of Regents’ finance and management reform proposal in response to the Court of Appeals ruling. The PJRCM approach is the basis for work performed by the American Institutes for Research and Management Analysis and Planning (AIR/MAP) for the Campaign for Fiscal Equity proposal. The CF approach has been the basis for measures of adequacy for New York performed by William Duncombe and John Yinger of the Maxwell School at Syracuse University.

Downes provides in-depth critique of the PJRCM and CF approaches, arguing that they are the two approaches most frequently employed and which have sound conceptual basis. According to Downes, PJRCM has several potential advantages including employing the expertise of educators to determine the cost of an adequacy standard and allowing experts to include aspects of producing education, both inputs and outputs that are not well measured and thus may escape the cost function approach. The experts can take account of research in forming their judgments about specific components, such as class size, and not merely rely on their own experiences. However, it is not without disadvantages. First, it is very time consuming. Additionally, it requires some means of adjusting benchmark costs to other districts. AIR/MAP used the Chambers’ regional cost method approach, which does not fully account for student differentials. Third, professional judgment panels never consider the cost of resources being proposed; thus from the professional judgment panel’s perspective there is no cost to additional expenditures. Fourth, experts are in fact educators who may suffer from a conflict of interest or limited range of experience in forming their judgments.

The CF approach typically employs administrative data to statistically estimate a cost function for a given level of student performance. It too has disadvantages. First it is complex and therefore often difficult to explain to non-experts. Second the process relies on measuring all the relevant components with a relatively high degree of accuracy, which is difficult to accomplish. Finally, the method does not describe how funds should be spent but rather only indicates how much is required. Downes concludes that the “best” approach may be to combine the cost function approach with aspects of professional judgment. For example, one could use the cost function approach to generate the right level of spending, and then require experts to allocate resources under the CF-established spending constraint.

Downes compares the distribution of resource allocation across several alternative plans: the AIR/MAP proposal (PJRCM); the Regents proposal; a teacher cost index proposal; a geographic cost index proposal; and a CF proposal by Duncombe. Downes finds that the AIR/MAP and geographic cost index proposals distribute a smaller portion of resources estimated for adequate spending to NYC than do the Regents, CF and teacher cost index proposals. The AIR/MAP and geographic cost index proposals would distribute about 40 percent of resources estimated for adequate spending to NYC, while the other approaches would distribute about 50 percent.

Sonstelie

In “[Financing Adequate Resources for New York Public Schools](#)” Jon Sonstelie examines alternative approaches to raising revenue for the local share of K-12 education expenditures. He examines five different revenue systems that would provide districts with revenues equal to \$13 billion, the level the NYS Regents propose as the local share needed to provide adequate expenditures. The five systems Sonstelie examines are: a) districts obtain general revenue through a statewide property tax; b) districts levy the same property tax rate as in a) above, supplemented by the state as needed to cover costs; c) districts levy the same rate, except for those that would generate revenues greater than costs; those districts cover costs only; d) each district’s tax rate is proportional to the income of its residents (the NYS Regents proposal); and e) the state and districts share the cost of the system inversely related to residents’ income (the system used in Maryland). These systems have very different implications.

For example, New York State could employ a statewide property tax system in which all districts levy the same property tax rate (e.g., statewide, 1.22 percent is required to raise the \$13 billion proposed by the NYS Regents). Such a system would require that 131 districts in New York raise more revenue than their costs, allowing the state to effectively recapture and redistribute that revenue for use in other districts. This occurs because property wealth is not distributed proportional to costs. Districts whose revenues exceed costs at a state uniform tax rate are more likely to have high values of residential property per household and low numbers of students per household relative to districts whose costs exceed revenues at the same tax rate.

Without recapture, a uniform state tax rate would need to be greater - 1.44 percent vs. 1.22 percent - in order to raise the required \$13 billion. Sonstelie estimates that approximately 139 relatively small, wealthy districts could levy tax rates less than 1.44 percent, and still raise sufficient revenues.

Under the NYS Regents proposal, districts are expected to levy a tax rate that is proportionate to their income per student, with state aid covering the difference between the revenue raised by this levy and expenditures needed to reach an adequacy threshold. Such a plan is very progressive. However, without recapture or a total expenditure requirement by districts, progressivity would be diluted. Adding to the dilution is a fair degree of variation in tax rates for districts with similar median household income, owing to the low correlation between household income and real property value per pupil.

Sonstelie then simulates a proposal similar to that recently implemented in Maryland. The Maryland proposal assumes a sharing relationship between the state and districts whereby each district receives the same share from the state, making up the difference between that level and adequate expenditures from local revenues. Such a plan is regressive as educational costs are inversely related to income, while property tax wealth is unrelated to costs. However, this plan does preserve local discretion.

Stiefel, Rubenstein and Schwartz

In "[From Districts to Schools: The Distribution of Resources Across Schools in Big City School Districts](#)", Leanna Stiefel, Ross Rubenstein and Amy Ellen Schwartz examine the determinants of resource disparities among schools in three large urban school districts: New York City, Cleveland, and Columbus, Ohio. In New York City, schools with a disproportionate share of poor and minority children typically have more teachers per pupil but those teachers tend to be less experienced and less educated than teachers in schools with relatively more affluent white children. The authors find wide variations in both school expenditures and the qualifications of teachers across schools. For example, schools with relatively high numbers of poor children have expenditures that exceed those of schools with fewer poor children. However, poor children tend to have teachers who are less likely to be licensed or hold a Masters degree. Similar results are found in analyzing school level expenditures in Cleveland and Columbus.

These three policy briefs on school finance have four important policy implications for education finance in New York State.

1. At this point there is no single best approach to measuring the cost to achieve an adequate education. Downes makes a good suggestion to combine the cost function approach with aspects of the professional judgment model. However, the method employed has important implications for the distribution of resources across schools and districts needed to achieve an adequate education.
2. Existing methods of generating and distributing resources rely on past relationships between spending and student performance. Student performance in turn relies on many factors other than expenditures. If these other factors substantially change, estimates of the cost of an adequate education will likely be misleading and the effects could be substantial.
3. Resource adequacy is largely a consequence of state standards and accountability is an important element of standards. This then brings into question the role of local discretion in raising revenues. Standards, especially rigorous ones, imply that the State should be convinced that schools have the resources they need to achieve the standards. If schools lack needed resources, it becomes very difficult to hold them accountable for student outcomes.
4. School finance and school accountability should be intimately linked. For example, if schools are to be the unit of analysis for accountability, schools in turn should be given discretion over budgets. In this case, resources could be tied to

student attributes so that schools with students in which the cost of achieving at least an adequate level of performance is greater receive resources to meet these needs. This has implications for how the State allocates aid.

Recruiting and Retaining Teachers Policy Briefs

Two policy briefs examine issues related to teaching and provide reinforcing findings and recommendations for policy. Don Boyd, Hamilton Lankford, and Susanna Loeb (BLL) offer a number of policy suggestions aimed at offsetting the concentration of low performance they document within New York State schools. Dale Ballou deals more broadly with compensation and human resource policies and their impact on teacher recruitment and retention.

Boyd, Lankford and Loeb

In “Improving Student Outcomes: The Role of Teacher Workforce Policies”, Boyd, Lankford and Loeb (BLL) provide a series of powerful statistics that demonstrate the remarkable and disturbing degree of concentration that characterizes the distribution of pupil performance within New York State. For example, they report that almost one half of all students who scored at the lowest level on the 4th grade mathematics examination in 2002 were located in just ten percent of New York’s schools. Further, they found that close to seventy percent of these students were concentrated in just twenty percent of the schools. The statistics for the highest performing students are also striking. BLL report that the majority of schools in New York State have very few students in the lowest performance group. Indeed, in 2002 28% of New York State schools had no students scoring at the lowest level. The authors report similar findings for fourth grade English Language Arts as well as for the eighth grade mathematics and English Language Arts examinations.

The authors also provide insight into the identity of these students and show that the vast majority attending the schools with the highest failure rates come from low income families and are either African-American or Hispanic. The authors also demonstrate the irrefutable fact that these schools with high failure rates are located in urban districts, with a great concentration in New York City.

Their analysis goes further and clearly shows the sizeable differences in the characteristics of the teachers teaching in these low performing schools. For example, BLL show that students in the schools with the highest failure rates are much more likely to be taught by an inexperienced and/or uncertified teacher with a poor record of performance on the basic liberal arts certification examination.

BLL view these findings with alarm given the converging evidence suggesting that the contributions teachers make to student learning are important and exceed the effects of other schooling inputs. They also make the sobering point that efforts in general to improve teacher recruitment and retention are not likely to improve the prospects for students and teachers in the lowest performing schools since there appear to be powerful forces which sort the teachers with the least impressive credentials in the direction of the students who are in the greatest need for high quality teaching. BLL conclude that these

adverse sorting effects must be addressed directly by policy makers and identify several options.

The authors emphasize the importance of compensation policy and note that research clearly shows that teachers are in fact sensitive to wage differentials. BLL caution against broad across the board salary increments in part because broad interventions like this would have no impact on the ability of failing schools in urban areas to compete more effectively with the suburban districts with whom they must compete. To bolster their points BLL present intriguing teacher turnover data that shows significant recent increases in the ability of suburban districts to attract teaching talent out of urban districts. In place of across the board increases, BLL encourage the use of more targeted programs designed to work to the benefit of both hard to staff schools and hard to staff subject areas. They also stress the importance of focusing attention on teachers early in their careers.

BLL encourage improvements in working conditions especially for teachers in hard to staff schools. The areas they identify for attention are: reducing class size; allowing more preparation time; improving facilities; improving the handling of disruptive students; and enhancing teacher induction and professional development programs. Increased attention to safety is a logical addition to this list.

BLL also identify a series of human resource policies they believe could have beneficial effects. One human resource problem they note is the tendency for urban districts to hire late in the hiring season. These late hiring practices are by-products of a complex set of interconnected policies and practices, including teacher transfer policies within urban districts. While it may not be easy to make the requisite changes, BLL make a strong case for trying to do so and note that these would not be costly reforms to initiate.

Finally, BLL encourage a continuation of what they refer to as “grow your own” strategies. They single out the Teaching Fellows Program in New York City as a promising example of using the desire of aspiring teachers to teach close to where they grew up. They also recognize the importance of attracting academically strong “home grown” talent into the schools.

Ballou

Dale Ballou’s policy brief “Improving the Teaching Workforce in New York Urban Schools” covers similar ground and provides a balanced assessment of the merits of several specific kinds of interventions, including reforms of compensation as well as reforms of broader human resource policies that are intended to improve recruitment and/or retention of teachers. Ballou also addresses the balance between state and local authority and proposes a number of novel means by which a state can intervene to offset powerful forces at the local level.

With respect to compensation, Ballou identifies a number of ways to differentially reward teachers. In particular, he examines premiums for new teachers such as lump sum signing bonuses, loan forgiveness programs, and mortgage assistance tied to a

commitment to stay in the district. He sees these as tools that should be targeted to recruit and retain high performing teachers who are willing to teach in shortage areas or in difficult to staff schools.

Ballou provides an insightful assessment of teacher merit pay plans in which he identifies steps that can be taken to enhance the viability of the plans. His suggestions include providing better training for evaluators and making use of an external fixed standard for the award with no cap on the number of possible winners. He also offers a list of precautions including: limiting the re-use of items on pupil assessments; testing to see if the results generalize beyond a single test; placing emphasis on long term rather than short term gains; broadening the range of outcomes that count; and including direct observation of teaching as part of the assessment.

With respect to the non-compensation areas, Ballou makes recommendations that are similar to those in BLL. He discusses the importance of making timely offers to teacher candidates and considers the role of disruptive students in making a teaching assignment attractive. He also raises an important question about whether it is reasonable to require a teacher to complete a subject area major before being eligible to teach. This is a controversial area where the debate is lively and where definitive evidence is in short supply (Wilson, Floden, and Ferrini-Mundy, 2001). Ballou makes the more general point that it ought to be possible for districts to recruit teachers from larger and non-traditional pools of available talent.

He also considers human resource policies that bear on the dismissal of teachers for cause and urges reforms that would make it easier for districts to remove teachers whose performance has been found to be deficient. This is another area of considerable controversy where the difficulties surrounding attempts to obtain clear, objective, and accurate assessment of teacher effectiveness have proven to be serious stumbling blocks.

With respect to governance issues, Ballou makes the argument that the state can and should intervene in ways that are designed to offset local decision makers. The example he uses involves the payment of premiums to certain types of teachers. He argues that such a policy will be resisted quite strenuously at the local level by teacher organizations and bargaining units and that the state would be wise to intervene by making the payments directly to the qualifying teachers. He imagines the local bargaining units will not be pleased by such direct payments, but he anticipates that their antipathy will not be so great as to prompt efforts to offset the state payments with local penalties, such as local reductions in salary.

He also proposes a novel state intervention to address whatever tendency there might be for suburban districts to recruit talented teachers away from their urban counterparts. He envisions a “robin hood tax” that would penalize suburban districts for “poaching” teachers out of urban districts. The tax or poaching fee would be imposed through adjustments in state aid and would be designed to discourage the practice and also to provide a revenue stream to the affected urban districts.

Ballou agrees with BLL about the absence of clear guidance about how large premiums need to be in order to influence behavior in the desired direction. He argues that one logical response would be to develop a bidding system that if enacted would force teachers to reveal information about their underlying preferences and thereby allow the state to avoid making the premiums unnecessarily large and thereby minimize costs.

Finally, he also offers caution about the potential for perverse consequences to emerge from the use of high stakes financial incentives. The example he cites involves National Board Certification and the rewards that have been attached by some states to National Board Certification for teachers. In addition to voicing a concern over the absence of definitive evidence linking Board Certification with actual pupil performance gains, Ballou notes the absence of security provisions that can raise questions about whether the submitted evidence in fact is coming from the person seeking certification. As the stakes become higher, incentives for applicants to tamper with the evidence will become greater.

We see five lessons that can be learned from the BLL and Ballou policy briefs.

1. There are glaring and disturbing levels of concentration of poor performance in the schools of New York State.
2. The State needs to focus its teacher recruitment and retention efforts on leveling the playing field between the schools at the lowest levels of performance and the others.
3. There is sufficient guidance in the research to justify developing targeted policies that are designed to improve the supply of teachers in shortage content areas to difficult to staff schools.
4. There are some novel approaches to targeted state intervention that are described in the policy briefs and that warrant further exploration.
5. Great care needs to be exercised to monitor results given the danger of perverse effects and the absence of clear insight into the required size of the interventions.

Incentives and Organizational Structure Policy Briefs

The incentives and organizational structure policy briefs provide an instructive contrast between an economic and more of a sociological/organizational behavior approach to matters of accountability. Helen Ladd develops the economic perspective and reports on North Carolina's recent experience with a comprehensive accountability structure. James Spillane raises a number of organizational issues that are not typically considered when economists think in terms of incentives and accountability.

Helen Ladd

In “Lessons from North Carolina's School Based Accountability Program” Ladd spells out the ingredients of North Carolina’s ABC program (Accountability, Basic Skills, and Local Control) and argues that North Carolina has made real progress but still faces significant challenges. In particular, she attributes the success of the North Carolina program to a combination of the legislature’s willingness to provide new funding and an approach to accountability that is positive rather than punitive.

At an intuitive level, the ABC program has considerable appeal. The basic idea is to reward teachers and staff in schools judged to be effective in raising student performance with financial bonuses. In contrast, schools judged to be low performing become the focus of state intervention. Ladd stresses the importance of measuring achievement gains rather than levels, despite the attendant complexity, and sees North Carolina as a place that has made progress toward developing the proper measures. She also notes that the North Carolina approach falls short of a pure value added approach. As a consequence, there is some remaining tendency for teachers of advantaged students to be rewarded not necessarily because of their merit as teachers but because of the collinear nature of pupil affluence and good test score performance.

Evaluations of the program suggest that it is having a positive impact on student achievement. North Carolina reports test score gains between 1993 and 2002 along with a drop in the share of low-performing schools. These gains have also been corroborated by increasing National Assessment of Educational Progress (NAEP) scores. Ladd notes that these findings are consistent with a recent study of accountability systems nationwide which found that strong accountability measures are related to performance improvements (Carnoy and Loeb, 2002).

One might be tempted to conclude from Ladd’s overview that strong accountability systems are highly successful and should build on the North Carolina (and Texas and Florida) successes. It is clear that these accountability systems are powerful and can make significant differences in how teachers and other actors within the system behave. It is also clear, however, that our understanding of these behavioral responses and their consequences for achievement gains is rudimentary at best.

Ladd points out a number of the more significant gaps in the knowledge base. She is sensitive to the potential for accountability systems to have unintended, and potentially harmful, effects. She conducted interviews with principals in North Carolina and reports their concerns about how the ABC program prompted significant shifts in resources away from non-reading and non-mathematics areas of the curriculum. She also notes the potential for an accountability system to make it more difficult for low performing schools to make improvements. In particular, she calls attention to the difficulties a low performing school may have retaining good teachers and expresses the concern that the accountability system can inadvertently contribute to inequality in the system. She also notes the sobering challenges that surround the measurement of performance. In addition to the already mentioned difficulties of achieving true value-added measures, she points out the challenges surrounding expectations for sub-group breakdowns of pupil performance required by No Child Left Behind legislation. Ladd questions the wisdom

of developing school-specific sub-group test score breakdowns and suggests that the district would be a much better organizational level for these metrics. She also considers the question of whether parental choice is a necessary ingredient of an accountability system and notes that the absence of a choice feature in the North Carolina plan did not obviously interfere with the success of the program.

Helen Ladd reaches the overall conclusion that accountability systems like the ABC program in North Carolina are promising and powerful, but recognizes that work is needed to fine-tune the design and to assess both the short term and long term effects. She attributes the success of the North Carolina program to the positive orientation of the system and to the success the state has had at fostering communication across the various levels of governance. Ladd's point about the importance of good communication leads directly to the James Spillane's policy brief in which he gives careful consideration to the potential for communication to break down across levels of governance in the pursuit of reform.

James Spillane

In "Policy in Practice: Where the Rubber Meets the Road" Spillane stresses the importance of the role human cognition ultimately plays in the implementation of policy. He argues that teachers, as well as others within schooling systems, see reforms "through the lens of their current practice," and that human sense-making tends to have conserving influences. In his view, there is a gravitas to current practice that generates resistance to change. The resulting coping mechanisms are impressive in both their creativity and effectiveness at preserving the status quo and in their success at blocking the intentions of reformers who make policy at more central levels of the system. Thus, for a centrally driven reform to work, like the ABC accountability system imposed by the North Carolina legislature, ways must be found to offset the naturally occurring conserving influences at less centralized levels of the system.

Spillane provides some compelling examples to illustrate his points. He depicts two teachers operating in the context of a common reform. The approaches employed by these teachers vary enormously, but in each case the response can be viewed as a good faith effort to be responsive to what the teacher understands to be the goal of the reform. The sharp difference in how these teachers respond provides powerful insight into why it can be so difficult to demonstrate the impact of reform efforts.

There are several interesting features to Spillane's argument. First, his approach contrasts rather strikingly with more conventional analyses of incentives and organizational structure, particularly those conducted by economists. Economists tend to focus on behavioral responses to incentives and sanctions and leave the analysis of thinking to others. Spillane takes the view that this narrowness misses some important insights.

Second, Spillane emphasizes a more benign explanation for the breakdown in communication that can occur across decision making levels in organizations. Spillane views the teachers in his example as making good faith efforts to interpret the reform. He attributes the difficulties that can occur in correctly interpreting the message to things like inattention, misperception, and reluctance to abandon pre-existing ideas, rather than to overt hostility and/or a fundamental difference in the goals being pursued at the school and teacher levels.

Third, this benign view of the source of communication difficulties leads to an optimistic view of what administrative leadership can do to achieve improved results. In particular, Spillane believes administrative leadership can play a helpful role in making relevant data more accessible and understandable, making the implementation process less solitary and more communal, and in being sensitive to needed variation in implementation practices across curricular content areas.

While it may be the case that the difficulties Spillane traces to human cognition can be overcome simply by having strong enough incentives and sanctions, it does not follow that this is the best approach to pursue. That is, why resort to aggressive incentives and sanctions if a kinder, gentler approach attentive to human cognition can suffice? An admonition to go as gently as possible is consistent with Ladd's points about needing to be cautious with accountability systems given their potential power and the very limited knowledge about their full effects.

While we are encouraged by the optimism that is part of Spillane's analysis, one nagging worry is that he may be overly generous in discounting the role that fundamentally different goals among actors at different decision making levels can and does play. Recall the earlier reference we made to Dale Ballou's observation that sometimes the state needs to intervene in ways designed explicitly to offset forces that are at work at local levels. It therefore seems prudent to at least recognize the possibility that decision makers at local levels can be impediments to reform. Of course, we can turn the matter around and also recognize that decision makers at the central level can also be impediments to grass-roots efforts at genuine reform.

Suffice it to say that hostilities can build up across the levels of decision making for any number of reasons, and presumably do not dissipate readily. The features Spillane suggests for accountability systems (ready access to relevant data, emphasis on communal rather than solitary implementation strategies, and sensitivity to necessary differences across curricular areas) are all prudent steps to take. Policy makers are also well advised to follow the North Carolina example of paying explicit attention to communication issues (thereby attending to the human cognition issues raised by Spillane) to minimize the severity of both the rewards and sanctions that are part of the system. Attentiveness to communication issues should be viewed as an investment that ultimately reduces the cost of the reform. Costly brute-force approaches to accountability can be avoided in favor of more nuanced approaches that are attentive to the underlying cognitive realities.

We see five lessons that can be learned from the Ladd and Spillane policy briefs.

1. Accountability systems have been shown to be powerful with beneficial effects, but are not a panacea. They require resources and good quality data, can easily get off track if communications break down, and are likely to have unintended consequences.
2. It is important to pay careful attention to technical details and as accurately as possible measure actual value added along outcome dimensions that are truly important to the state.
3. It is also important to pay careful attention to human cognition issues since success here can reduce the magnitude of the rewards and sanctions that are part of the accountability system. In particular, efforts are warranted to make relevant data available and understandable, to make implementation processes include teachers and administrators, and to allow for differences across curricular areas.
4. In addition to the benign reasons for communication failures, attention should be paid to overtly as well as covertly hostile actors in the system who are pursuing agendas that run counter to the goals of the reform.
5. Constant monitoring and careful periodic evaluation are important ingredients for success. The state is well advised to monitor accountability systems carefully and to take corrective steps along the way as more knowledge is gained about their actual impact.

Data Use and Management Policy Briefs

Richard Murnane and Nancy Sharkey examine how the results of student assessments can be used to improve student outcomes in New York City. They employ a case study of the New York City Department of Education's experience with two consultants, the Grow Network and the Princeton Review, to use student assessments to alter what happens in schools and classrooms. Anthony Cresswell and Sharon Dawes examine the information related challenges posed by the Campaign for Fiscal Equity court decision and propose some possible solutions.

Murnane and Sharkey

Murnane and Sharkey, in "[Learning from Student Assessment Results: Opportunities and Obstacles](#)" find that assessments are typically used for three purposes: 1) to make decisions regarding student academic paths, e.g., summer school or grade retention; 2) for formative or summative judgments regarding what students know, what they are able to do, or the effectiveness of teaching; and 3) symbolic support for predetermined positions.

The authors find that the use of student assessments may provide important information to effectively they alter teacher behavior and improve student test performance. Murnane and Sharkey argue that student assessments are much more likely to have the intended effects when: a) assessments are clearly aligned to state standards and to the curriculum being taught; b) the technology available to analyze test items is user friendly for teachers with modest levels of computer literacy; c) teachers know how to analyze test results and derive implications for changes in instruction; d) teachers have time, ideally in the school day, to do this work; and e) the school culture embraces the ideas that the achievement of all students is the responsibility of the entire teaching staff and that student success depends on continued learning.

Murnane and Sharkey offer several specific policy recommendations based on their analysis of New York City's use of student assessments to improve student learning and teaching: 1) New York State should provide school-based educators with the results of student assessments as rapidly as possible. 2) The state should contract for the development of formative assessments linked to standards for local districts to use. 3) The state should push forward as rapidly as possible with its plan to develop a state-wide student identification system. 4) The state should examine whether score trends on its and NYC's high stakes tests are consistent with trends in the scores of New York State and NYC students on the National Assessment of Educational Progress (NAEP). 5) The state should examine the effectiveness of the system under which Boards of Cooperative Education Services (BOCES) provide New York school districts with software and training for making constructive use of student assessment results.

Cresswell and Dawes

In their brief "[The Information Dimension of Education Financing Decisions: Data Needs, Users, Strategies, and Systems](#)" Anthony Cresswell and Sharon Dawes examine the information-related challenges posed by the Campaign for Fiscal Equity (CFE) court decision and propose some possible solutions. Since CFE focuses primarily on student outcomes, Cresswell and Dawes argue that the primary information challenge is to accurately measure student performance on an array of outcomes that reflect the broad goals of education, followed by determining how to measure the use of resources required to improve these outcomes.

The authors focus on two unresolved questions: 1. How do we accurately define and measure "sound basic education"? and 2. Is the current information collection and reporting system up to the task? The authors express schooling in terms of *inputs* (resources and decisions that go into a system), *process* (how the school environment uses the resources), *outputs* (achievement, graduation rates, etc.) and *outcomes* (post-schooling results, as in community contribution). The CFE decision focuses on the post-schooling outcomes (as in "meaningful civic participation") which the authors argue lack measurement systems, and therefore place greater emphasis on reviewing the *processes* that promote academic achievement and civic participation. They argue that because of this shift in focus away from outcomes, the current structure and content of the information systems is inadequate as a measurement tool. A major deficiency in the current information system is its inability to document the process of learning needed to

understand how and why students fail to reach standards. As a result, systematic improvements are unlikely to occur.

The authors present an overview of the current information frameworks available to and used by school districts (e.g., LEAP, STEP and BEDS) which they state are adequate for providing snapshot quantitative analyses, such as what is happening where, and comparisons. However, the authors contend they are not useful for explaining why certain patterns occur, or supporting policy changes that might provide sustainable systemic improvements, which they argue is what the CFE charge calls for.

The authors propose five approaches to improving the information environment in education: 1) Investments in statewide single-purpose information systems; 2) data warehouses and associated decision support systems; 3) systems integration projects; 4) data quality initiatives and 5) enterprise information architecture initiatives. Dawes and Cresswell view the enterprise initiative as the one providing the greatest promise. Enterprise initiatives re-conceptualize K-12 education as a holistic enterprise, with interdependent goals, structures, processes, actors, information content, and information systems. Their interdependent relationships need to be understood and accounted for in the development of strategies, programs, and systems.

The Murnane and Sharkey and Cresswell and Dawes briefs have several, meaningful implications for policy.

1. It is important to have information systems that translate student achievement results into useful information for teachers and principals so that they can alter their behaviors to improve student outcomes. There is some evidence that the current system is unable to provide the requisite information.
2. The development of formative assessments linked to the learning standards, which can provide quick feedback to teachers and principals during the year, would substantially improve the accountability linkage between the learning standards and teachers.
3. Current testing is an important and useful initial signal about performance in the system. However, this approach can evolve to viewing the K-12 system as a holistic set of interrelated and interdependent processes with desired outcomes of which tests may be only a weak signal. The evolution of this process will increase the linkage between the structure of education, teaching practice, and student outcomes. Testing will continue to play an important role but can be refined to more clearly reflect the connection between schooling and desirable student outcomes.

III. Conclusions and Implications for Policy

Taken together these observations provide useful insights for policy. They also highlight some longstanding dilemmas. Several of the authors indicate the importance of

differences between schools within large districts. Others recognize the difficulty of defining problems so narrowly that resource disparities make equitable solutions difficult. New York State has often focused on the district as the unit of financial accountability but increasingly the school is the unit of performance accountability. Is this a misalignment that contributes to the difficulty of increasing student performance? If so, how can it be resolved?

The other tension that exists in some of the policy briefs is that between affording local administrators the discretion and authority to determine how to address performance problems and the value of having the state apply a common set of requirements and policies across all districts. For example, everyone recognizes the importance of increasing the quality of teaching, especially in low-performing schools. Should superintendents and principals have the discretion to implement policies that take account of the local environment and only be held accountable for the student performance outcomes? Alternatively, should the State impose uniform education and certification requirements? The right balance of organizational structure and incentives remains an important issue.

While these questions are not resolved in the policy briefs, the briefs do highlight some important decisions that policy makers should consider regarding education finance and accountability. Finally, data collection, management and dissemination can play an important role in facilitating improved student performance. These are topics where we believe more research is needed.

References

- Carnoy, Martin and Susanna Loeb (2003). "Does External Accountability Affect Student Outcomes? A Cross State Analysis," *Education Evaluation and Policy Analysis*, 24(4). 305-331.
- Wilson, Suzanne M., Robert E. Floden, and Joan Ferrini-Mundy (2001). *Teacher Preparation Research: Current Knowledge, Gaps, and Recommendations*. A research report prepared for the U.S. Department of Education by the Center for the Study of Teaching and Policy, University of Washington.

Appendix A

Symposium on Education Finance and Organizational Structure in New York State Schools

Education Finance Panel

Thomas Downes, *Tufts University*

“What Is Adequate? Operationalizing the Concept of Adequacy for New York”

Jon Sonstelie, *University of California – Santa Barbara*

“Financing Adequate Resources for New York Public Schools”

Leanna Stiefel, Amy Schwartz, *New York University* & Ross Rubenstein, *Syracuse University*

“From Districts to Schools: The Distribution of Resources Across Schools in Big City School Districts”

Teaching Workforce Panel

Donald Boyd, Hamilton Lankford, *University at Albany*, and Susanna Loeb, *Stanford University*

“Improving Student Outcomes: The Role of Teacher Workforce Policies”

Dale Ballou, *Vanderbilt University*

“Improving the Teaching Workforce in New York Urban Schools”

Organizational Structure Panel

Helen F. Ladd, *Duke University*

“Lessons from North Carolina's School Based Accountability Program”

Jim Spillane, *Northwestern University*

“Policy in Practice: Where the Rubber Meets the Road”

Data Management Panel

Anthony Cresswell and Sharon S. Dawes, *University at Albany*

“The Information Dimension of Education Financing Decisions: Data Needs, Users, Strategies, and Systems”

Richard Murnane and Nancy Sharkey, *Harvard University*

“Learning from Student Assessment Results: Opportunities and Obstacles”