

Promoting Accountability In New York State

A Working Paper
Prepared as Support for
“Can New York Get an A
in School Finance Reform?”
A Report by
the Citizens Budget
Commission
November 2004



FOREWORD

Founded in 1932, the Citizens Budget Commission (CBC) is a nonprofit, nonpartisan civic organization devoted to influencing constructive change in the finances and services of New York State and New York City governments.

This Working Paper was prepared under the auspices of CBC's Education Finance Committee, which we co-chair. The other members of the Committee are: Paul R. Alter, Richard H. Bagger, Stephen Berger, Deborah A. Buresh, Lawrence B. Bittenwieser, Evan A. Davis, Stephen F. DeGroat, Kenneth D. Gibbs, Bud H. Gibbs, H. Dale Hemmerdinger, Bill Lambert, James L. Lipscomb, Stanley Litow, Robinson Markel, Joel H. Moser, David I. Moskowitz, Lester Pollack, Hector P. Prud'homme, Alfredo S. Quintero, Edward L. Sadowsky, William G. Salter, and Howard Wilson.

The CBC created this ad hoc Committee in the wake of the 2003 New York State Court of Appeals decision in the Campaign for Fiscal equity case. Initially, we believed that the scope of the Committee's work would be confined to helping public officials identify the most economically effective sources of revenue to fund the Court's decision. However, as the Committee explored the issues, and as the Governor and State Legislature failed to reach an agreement on the amount required, we broadened the scope of our work. The Committee recognized that in order for the goal of the Court's ruling to be achieved - for students to obtain a sound basic education - changes were needed beyond the allocation of more money.

The Committee prepared its final report, *Can New York Get An "A" In School Finance Reform*, with two goals in mind. The first is to provide the responsible public officials – judges, legislators and others – with sound advice on how to craft a remedy that will be effective and efficient. But the CBC also recognizes that shaping policy affecting so many lives, and costing so many billions of dollars, should involve an informed citizenry who support the eventual outcome. Accordingly, the CBC seeks also to use the Committee's final report, and a companion conference scheduled for December 2-3, 2004, to stimulate informed debate about the options available to New Yorkers for providing their children a sound basic education.

In order to prepare the final report, the Committee met nine times between January and November of 2004. Its meetings were focused on research conducted by the staff and expert consultants. This background research has been organized into six working papers, of which this is one. This working paper was prepared with the help of Anthony Shorris, who teaches at Princeton University's Woodrow Wilson School of Public and International Affairs and directs the School's Policy Research Center for the Region. He served as a consultant during the Committee's early meetings. All six working papers and the final report are available at the Commission's website, www.cbcny.org.

The research was made possible by generous support from the Andrew W. Mellon Foundation and by designated contributions from CBC Trustees.

Paul E. Francis
Eugene J. Keilin

PROMOTING ACCOUNTABILITY IN NEW YORK STATE

Any significant increase in the level of resources available for public education in New York State has the potential to be wasted without effective programs for its use. While better instruction alone can help increase student performance, the dramatic increases called for by the Court are only possible with both an infusion of resources *and* real reform in the delivery of educational services in much of the State.

In examining how money – new money as well as the money currently spent by New York’s districts – is used, the State needs to develop better accountability systems than are now in place. The current accountability systems have grown up over the years and are inadequate and sometimes even conflicting (as described below, a single school in New York City may find its principal receiving a performance bonus while the school is listed as “in need of improvement” in the federal accountability system and “under registration review” in State terms). Alignment of the three tiers of accountability systems by which schools are measured is a pre-requisite to any effectiveness; conflicting accountability systems are a guarantee of ineffectiveness.

One important point must be kept in mind when considering the implementation of new accountability systems. Change in education always comes slowly. No amount of intervention, no matter how effective, can make up in a year what has been lost over many years of a child’s education. Teachers must be hired, trained, and re-trained; curricula must be selected, implemented, and modified; school leadership must be trained or replaced; accountability measures must be developed, implemented, and consequences felt. When the No Child Left Behind Act (NCLB) was passed, the federal government gave states 12 years to meet its extremely ambitious goals of having every student meet State standards. While some may quarrel with other aspects of the NCLB law, this recognition of the timeframe required for real reform is certainly appropriate and something New York must keep in mind. No matter how large the increase in resources may be – and to some extent the greater the increase the longer it will take it be absorbed – significant and demonstrable change in student performance will take a number of years to be realized.

This lengthy timeframe makes it even more important that the State establish effective accountability systems before significant incremental money begins to flow so that New Yorkers will have a baseline against which to evaluate the efficacy of new funding. Further, any accountability system will require fine-tuning over time as school leaders adjust to it and the effects of the positive – and any unplanned negative – incentives it creates are felt.

Some efforts have been made to describe better accountability systems for the State, including proposals made by the CFE and the Zarb Commission. These proposals focus on better planning for the use of the new funds, but do not go far enough in laying out clear, consistent, and measurable goals for the State’s educational system.

Accountability: Two Dimensions

At least two issues are central to effective accountability systems: how the resources are used and whether the outcomes from the use of those are adequate. A new accountability system needs to include both perspectives.

1. Input Accountability: Using Money for Purposes Intended

The State Court of Appeals has specified the key focuses of increased State spending on education, and the State's accountability system should ensure that the enormous infusion of funds associate with the CFE case are targeted to meet the requirements of the Court. Indeed, the very language of the decision specifies such an accountability system. It is important that such an effort not be entirely retrospective, but include a requirement that each district present a plan for the utilization of any new funds it receives in accord with the mandates of the Court. Encouraging the leveraging of existing funding to support new instructional programs and more effective management should also be a part of the accountability system. Reporting on these efforts should replace or at least be integrated with other State reporting requirements to avoid duplication and additional administrative costs. Finally, no funds should be allocated to purposes other than those specified by the Court and in the State response to it without the prior consent of the State Education Department.

2. Outcome Accountability: Getting the Desired Educational Results Efficiently

Even if funds are dedicated to the purposes specified by the Court, they can certainly be used ineffectively or inefficiently without clear standards. CBC believes that clear, comprehensive, coherent, and statistically valid measures can be useful tools in ensuring the State's education dollars are used wisely. The measures of performance for the schools, districts and the State must reflect best practices in educational accountability systems, not simply reliance on the lowest cost tool – standardized testing. In addition to measures of effectiveness, the enormity of the additional funding called for by the Court requires that the accountability system also examine the efficiency of the delivery system in various districts and perhaps even schools.

Current Accountability Mechanisms: Federal, State, and Local

Schools in NYS currently operate under at least three different accountability systems – federal, State and local – each focused primarily on output rather than input, and none of which are well-aligned with the others.

Input accountability systems

Federal requirements relating to the uses of grant money are typically quite broad. While some categories of programs set specific guidelines, the larger programs such as those for special education and school lunch programs are not highly limiting. The plans and reports that federal grants require are generally *pro forma*, handled by grants management staffs in districts and sometimes large schools. The No Child Left Behind Act did not substantially change this approach – and focused on outcome measures. In a few cases, federal reporting requirements are incorporated into other documents (see below).

Input accountability at the State level is focused on the development of two levels of planning documents, the school-based Comprehensive Educational Plan (CEP) and the new district-based Regional District Comprehensive Educational Plan (RDCEP). These documents are designed to be the core planning documents around which the schools and districts are to build consensus on goals and strategies for the year to come.

- In the case of the individual schools, the CEPs are developed and approved by a committee composed of parents, teachers, and school leadership (designated as the School Leadership Team or SLT under an earlier State governance law and State regulations) that is charged with effecting school-based management. These documents are designed to include a needs assessment and statement of annual goals and measurable objectives, to highlight specific strategies to be used, to lay out an action plan, to develop a budget and to include some evaluation of the prior year.
- For districts, the RDCEPs, which in the case of New York City had been the product of the 32 community school boards and are now produced by the ten Regions on behalf of the 32 districts, also provide a single statement of the mission and approaches of the district to dealing with issues faced by the schools. In general terms, they follow the same structure as the CEPs, but on a district-wide basis. The documents can be very voluminous (in the past, DCEPs for some districts were more than 200 pages long) since they also serve as the key planning and reporting document for several different federal and state programs. However, by their very length they are often unusable as accountability tools, too detailed for any but the most involved parents and community leaders to use. While the form for the documents does include a linkage to the use of resources by the school and district, the programmatic relationship is hard to discern and not effective as input accountability.

Output accountability systems: schools

Output accountability tools are much more widespread – perhaps too much, so since federal, State, and local systems exist simultaneously. At the federal level, all schools in the United States that receive federal funds under Title I of the Elementary and Secondary Education Act (ESEA) as amended by the No Child Left Behind Act (NCLB) are subject the accountability standards set out by the federal government. Federal regulations require states to establish performance standards for grades 3-8 that must be achieved by virtually all students by 2012.

Schools are then measured by the annual yearly progress of students in the school – in total and for each of several demographic groups – towards the achievement of those standards. By 2005, all students in grades 3-8 are to be tested annually (something that New York City has been doing for many years, but other districts in the State have not, relying instead on State-developed tests in grades 4 and 8 as well as high school Regents exams). Schools failing to meet these progress goals are subject to increasing levels of sanction under federal law over a period of five years:

- After two years in which a school fails to meet its annual progress goals for the school as a whole or for any demographic group, it is identified as a “school in need of improvement”;
- In the second year in which progress is deemed inadequate, schools must allow students in them the right to transfer to other schools not in need of improvement, with transportation costs to be funded from existing Title I funds;
- In the third year in which a schools fails to make adequate yearly progress for any demographic group, they must designate an additional portion of their federal Title I funds for outside tutoring, to be provided by private vendors, the district, or other non-failing schools;
- In the fourth year, the school must begin to take corrective actions. These corrective actions must include at least one of the following: replacing school staff, implementing a new curriculum (with appropriate professional development), decreasing management authority at the school level, appointing an outside expert to advise the school, extending the school day or year or reorganizing the school internally; and,
- In the fifth year of inadequate progress, the school district must initiate plans to fundamentally restructure the school. This restructuring may include reopening the school as a charter school, replacing all or most of the school staff who are relevant to the failure to make adequate progress, or turning over school operations either to the state or to a private company with a demonstrated record of effectiveness the school must be closed, have its management outsourced, or be taken over by the State.

In a state such as New York with very high standards, these accountability measures are challenging. Indeed, as many as 40% of all schools in New York City have already been designated as “schools in need of improvement” and are on the road to suffering other federally-mandated sanctions.

New York State had already put its own accountability system for schools into place many years ago, one that presaged many of the federal NCLB requirements. Under pre-NCLB State regulations, schools failing to show progress towards meeting standards (using a different methodology than the federal rules) are subject to increasing sanctions, including being designated as a “school under registration review” (SURR), re-structuring, or closure. The State’s SURR system remains in place, paralleling but not identical to the federal system it is

also required to administer. Since State tests are currently administered only in grades 4 and 8, all State accountability systems for elementary and middle schools are based on just these two grades (with Regents exams used for high schools). Once the State begins to administer annual tests for all grades 3-8 as required by federal law by 2006, it will be possible to merge the federal and State systems, or at least to use a common basis for assessing performance rather than maintaining two inconsistent systems.

New York City has developed several accountability systems of its own. Given its annual testing process, all City accountability measures are based on the more sophisticated model of tracking changes in individual student performance year over prior year, aggregated up by grade, by school, and by district. This is commonly referred to as a “value-added” (or longitudinal) approach that is recommended by most educators. (By contrast, since the State only requires testing in the 4th and 8th grades, its accountability systems must be based on comparing performance of students in one grade of a school to a different group of students in the same grade the prior year, sometimes called a “cross-sectional” approach.) These mechanisms are in addition to the regular managerial evaluation program performed by superintendents for school leaders.

For several years, the City reported school and district instructional progress using District Performance Profiles (DPPs), tracking and publicly reporting growth in students’ tests scores from year to year by district, by school, and by grade in each school. These reports have apparently not been updated in nearly two years, or at least they have not been made public as they had in the past, and there appears to be no public district- or school-based reporting of scale-score progress as compared to similarly situated schools.

In addition, school principals and other school-based administrators have been subject to another accountability system through their Council of Supervisors and Administrators (CSA) contract with the City. This contract offers them cash rewards for strong performance. In the case of the CSA bonuses, the measurement used is the value-added assessment, with schools compared to other schools with similar demographics based on the year-over-year growth of children on standardized tests. Leadership in schools in the top cohorts (based on high growth rates among demographically-similar schools) is eligible for cash awards upon completion of a project to disseminate to other schools what tools they used to achieve the strong performance. Finally, some districts in the State may have their own measures of accountability for their schools, while others simply rely on the State measures, limited as they will be until the advent of the federally mandated annual testing required within two years.

Output accountability: districts

All of these accountability tools are focused on individual schools. Districts are rarely held accountable through such formal mechanisms, although federal law does require states to designate districts as “in need of improvement” based on the performance of their schools. If a district fails to make adequate yearly progress for straight four years, under federal law, the State must take corrective actions that must include at least one of the following: deferring programmatic funds or reducing administrative funds; implementing a new curriculum (with professional development); replacing personnel; establishing alternative governance

arrangements; appointing a receiver or trustee to administer the district in place of the superintendent and school board; or abolishing or restructuring the school district. The state may also authorize students to transfer, with transportation, to higher-performing public schools operated by another school district. States must provide information to parents and the public on any corrective action the state takes with school districts.

In New York City, the schools remain legally organized into 32 community school districts, and public presentations must be made using these units, but the districts have been essentially relegated to reporting entities, with all administrative authority for instructional purposes held in ten instructional Regions, each headed by a Regional Superintendent. (Operating authority has been strongly centralized, carried on through six regional operations centers.) The City will be reporting at this level as well, and the RDCEPs will be prepared by staff in the ten regional offices.

Finally, and perhaps most significantly, political accountability varies across the State. In the case of more than 99% of the districts in the State, ultimate accountability for performance of the schools in a district comes through the election of the school board by the voters in the community. The elected school boards bear the final accountability for the selection and retention of the superintendent and use of resources (the level of which may be separately voted upon by the electorate) that the schools receive. For the handful of other, so-called dependent districts outside NYC, funding decisions are made by the mayor and local city council but operational control remains in the hands of the elected school boards, a source of complex and sometimes confused accountability. And in the case of NYC, since 2002 the Mayor has been given complete control of the school system, with the essentially powerless new Panel on Educational Policy (nominally replacing the old Board of Education) serving at his pleasure and subject to his removal should they disagree with his policy direction. With the Chancellor now accountable only to the Mayor, meaningful accountability for improved public education in the State's largest jurisdiction is now exercised by its voters once every four years.

Problems with the Current System

The combination of various sets of accountability tools is immensely confusing for teachers, parents, and school leaders – the very people whose behavior they are designed to affect. The differing systems make it possible for a school to be at once outstanding and failing, improving and declining. This is especially true when the State and federal standards must be based on tests that are administered only in two grades and which compare the progress of students towards achievement of State standards of proficiency, while the City has opted to use systems based on the individual student growth. For example, a school with poor growth in its 4th grade results in one year as compared to the prior year could be showing inadequate yearly progress towards meeting standards on State and federal measures – and thus be designated a “school in need of improvement” in federal terms and a SURR school in State terms – and yet have its principal win a bonus through the principal's bonus provision of the CSA contract based on test score growth of students in other grades. This process became so counter-intuitive that the City decided to add a criterion to its principal bonus plan excluding principals of SURR schools,

even when they might otherwise be eligible. No accountability system can be successful if it sends conflicting signals about what is important and how it is to be measured.

A more significant issue in the State and even federal accountability system is the reliance on standardized tests. While there is little question such measures can be useful indicators of performance, especially at the extremes as opposed to the margins, there is also a vast array of research in the field indicating that inappropriate use of such tests risks distorting their purpose, and worse, potentially distorting the educational process. First among the many issues with the use of standardized tests is that they are not simply statistically reliable at the individual classroom level, often not reliable at the demographic group level within a school, and sometimes not reliable even at the school level, depending on the number of children taking the test. For example, an average K-5 school of 600 children will have 100 children in each grade. Under federal law and State standards, if more than 40 of these children fall into one demographic group (based on their race or their level of English proficiency) their results must be disaggregated and reported separately. If any one such group of students in a school fails to make annual year progress targets, the entire school is designated a “school in need of improvement” and all federal sanctions must be applied. By using a group as small as 40 children, as few as two or three children having a bad day can affect the future of the entire school.

More subtle issues also arise. Putting this much weight on a single standardized test makes an increasing focus on the tests themselves more likely. Subjects not tested are likely to be undervalued (indeed some NYC schools have already reported reductions in arts, social studies and elective programming so there can be more focus on tested materials). Further, such over-weighting of a single exam encourages teachers to focus on the material tested in the exam rather than a broader body of knowledge. More importantly, such tests encourage a kind of triage by teachers reacting to the implied incentives, spending their limited time on children near the margin and reducing the amount of time spent with the strongest and weakest students. This phenomenon is exacerbated by accountability mechanisms such as the State’s, which use the number of children meeting the standard, rather than year-over-year score changes for individual children, as the accountability tool – ensuring that teachers focus on children near the “cut” score. Pressure from such tests may encourage schools to minimize the number of potentially poor performers taking the test, by offering to label children as in need of special education (already reported in some New York City schools), or as having limited English proficiency (which gains children additional time on tests). Finally, of course, cheating by teachers and administrators, while not widespread, has been reported in schools and increased focus on the tests makes this only more likely.

The issue is more complex in high schools, where New York State’s choice to establish the highest standards in the nation for graduation – passing five separate Regents exams with a score of 65 or better – has led to growing evidence that increasing numbers of children are dropping out (even as others are striving harder than ever to meet the new standards). Statistically, the more exams that are required for graduation, the greater the odds that a student will perform below his or her actual skill level on one test, with high-stakes consequences. Issues in test design and implementation – such as the State’s problems last year in its math Regents

exams and continuing problems with the physics exams – become far too important when no other criteria can be used for graduation.

In addition, there is a commonly held assumption that standardized tests are by definition more objective than other forms of assessment. In fact, tests are developed based on judgment (groups of teachers are asked to estimate the relative difficulty of questions and the grade level of knowledge they represent); they are scored (in the case of the written elements of reading tests and all Regents exams) using judgment, and their application of the tests to promotion and other decisions often involves judgment. The objective elements of the standardized tests are important, and are useful but not without limitation.

Further, the accountability system makes no effort to examine efficiency as well as effectiveness, a critical gap when large increments of additional dollars may be directed at education in the State. All measures look to the change in performance, not the level of resources applied to achieve that change. While no financial analyst can quarrel with the need for dramatic improvement in student performance, no educator should quarrel with the notion that achieving such change in the most cost effective manner is desirable. Costs per student are already reported by school, and uses of funds are also included in reports issued by districts. What is not readily available to parents are appropriate comparative data, and some measure of the change in efficiency of their child's school in the current year as compared to prior years. For example, while school report cards already report spending per student, a better indication of efficiency might be a ratio of spending per student in a given grade to average scale score growth for the grade at this school as compared to this school the prior year and perhaps even to other schools with comparable demographics. Although absolute standards do not exist, relative changes in the same school over time and as compared to similar schools might be a step in the right direction.

Lastly, there are few positive consequences for good performance, and even the existing negative consequences of failing to meet accountability goals are often obscure, conflicting, and not enforced in a timely manner. Accountability without consequence is hardly accountability at all.

Currently, the only significant positive incentive available is the principal performance bonus in the NYC CSA contract. While a good first step, this program has been troubled over the years by the extended lag time in making the awards (they are not made until as much as a year after the conclusion of the school year), the complexity of the mechanism (which many principals report they do not understand), and worst of all, the impact of the bonuses on principals' pensions and retirement planning. With most principal pensions based on their wages in the final year of employment, the current system gives the best performing principals an incentive to retire immediately after winning a bonus – hardly the right signal.

The Chancellor has indicated that he will be considering some process of recognizing high performing schools in the coming year, perhaps offering such schools additional instructional or budgetary flexibility. This would be a powerful signal to good performing principals in a system that severely limits their flexibility.

Negative consequences are more real, although the school system in New York City remains unable to remove significant numbers of poor performing teachers or principals. While many contractually-defined rules make removals for poor performance more difficult to execute than they should be – and lead to a counter-productive practice of pushing poor performers out of one school into another rather than terminating them – the most important limitation on large-scale termination of poor performing instructors and school leaders is the paucity of quality on the “bench” behind them. With rates of turnover among principals and teachers already high due to retirement and recruitment to more attractive positions elsewhere, troubled schools and districts are in no position to remove many poor performers.

Negative consequences are also not accompanied by additional support. At the federal level, poor performing schools actually lose money (as NCLB requires their Title I money be re-directed to busing kids to other schools or paying for private tutors). State regulations that lead to sanctions offer little in the way of additional support, although programs such as the now-abandoned Chancellor’s District (see below) did focus additional resources of troubled schools.

Finally, negative sanctions are not likely to be effective if they are based on standards that cannot be achieved. NCLB requires that all students reach state standards by 2012, even where such standards are relatively high, as in New York State (some other states have already reduced their standards to accommodate this requirement). No one benefits by “dumbing down” State standards, or having different goals for different children when every parent wants their child to achieve a sound basic education, but using the ceiling (where all students meet all standards) as the floor (below which sanctions must be applied) is counter-productive. Having 100% of students meet high standards is a laudable goal, but there are few other fields where such an ambitious goal has been stated, and fewer still where sanctions are applied for failure to meet it. Everyone would support reduced rates of infant mortality, yet it is unlikely that any level of government would consider sanctioning health departments that fail to eliminate it entirely. New York State Regents standards that preclude diplomas without high passing scores in five subjects are also excellent and important objectives – but one may question whether they are realistic in the timeframe set out by the federal and State governments. Any system that leads to as many as half or even two-thirds of all schools being designated as “failing” and subject to closure is suspect. Such an approach is not realistic and risks engendering cynicism and a loss of confidence in public education rather than promoting the pursuit of excellence.

RECOMMENDATIONS FOR AN ENHANCED ACCOUNTABILITY SYSTEM FOR NEW YORK STATE

The first and most important general recommendation is that there should be clear alignment among the various accountability systems now operating across the State.

- ***Implement a consistent State-wide accountability program.*** CBC recommends that the State develop a single set of accountability tools that are uniform across the State, that are consistent with federal guidelines, and that replace locally-generated systems.

Such a system can build upon New York State's existing system (one of the first approved by the federal government), but must become more coherent, comprehensive, and sophisticated.

Input Accountability Recommendations

The first step in better management of the instructional delivery system of the state is to ensure that all funds – but most especially the new funds specified as part of the Court mandate – are used for purposes consistent with the requirements of the CFE case.

- ***New funds should only be used for purposes consistent with the CFE settlement.*** All districts should be required to specify, in advance of receipt of new funds, the uses to which the new money will be put. In order to reflect the specific needs of the community's students, those uses to which the new money is put should be selected by the local district from a listing of options provided by the State that are themselves reflective of the State's response to the Court. Any district seeking to use CFE funds for other purposes should be required to seek the specific approval of the State Education Department. Districts should then be required to report to the State annually on the actual uses of the new funds against the specified purposes as part of a comprehensive planning process that includes the uses to which all funds are put.

However, the State accountability system should not stop at the use of new monies, but should also attempt to direct existing funds to meet the Court mandate.

- ***Additional monies should be a catalyst for making the use of other funds more effective.*** In their District Comprehensive Educational Plans (DCEPs), filed with the State annually, all districts should be required to specify how other, *non-CFE funding* will be re-directed to the achievement of the goals laid out by the Court. Further, each District should be required to specify what other actions it will be taking to address the issues raised in the case – what curriculum changes, management improvements, and re-direction of other funds will be part of the District's plan for the year to come – and what specific outcomes parents can expect to see (in terms of reduced class sizes, new books, higher percentages of better trained teachers). These changes in other inputs to the delivery of educational services should be specified. No district with inadequate results should use the additional money for more of the same – CFE funds should be a catalyst for change in entire instructional delivery system.

Districts are generally resistant to specifying actual levels of improvement for fear of embarrassment if they should fail to meet the targets, but in this case, given the enormous investment of the State in this program, the risk is worth taking.

- ***Districts should be required to specify performance goals for the year.*** Each District should be required to specify what progress it will achieve in the next four years with the additional funds. Improvement should not only be in standardized test scores, but should include changes in graduation rates, college readiness, attendance, school violence, and other qualitative measures of school performance. This way districts

can be held truly accountable for the uses of the funds according to plans they have specified.

Outcome Accountability Recommendations: Effectiveness

There remains great controversy in the field concerning the effectiveness of standardized tests in assessing student learning. Assessing the appropriateness of such tests as pedagogical tools or their use for “high stakes” purposes is a matter better left to instructional experts, but for accountability purposes standardized tests should certainly be a part of the assessment program. However, they should be one aspect of a comprehensive program to ascertain whether the State’s new resources are having a meaningful effect on the State’s children. Otherwise, the tests themselves become the entire focus of the instructional effort, rather than one measure of its effectiveness. This program should include quantified measures as well as qualitative assessments of school and district performance, much as hospitals are assessed using both tools for purposes of their accreditation.

Quantitative evaluation: standardized tests

In the field of education accountability measurement, a growing consensus has come to view “value-added” assessments as the most meaningful. This approach compares a given student’s performance on standardized tests in a given year to the student’s own performance the prior year. By looking at individual children and comparing them only to themselves in prior years, this approach avoids issues of demographic differences and changes in class composition over time since children are not being compared to each other. For such tools to be effective, however, children need to be tested at least annually, something that New York City has been doing for many years in grades 3-8, and which all districts in the State will be required to do within two years under NCLB.

- ***Where standardized tests are used for accountability purposes, they should be based on “value-added” methodologies.*** Once an annual testing protocol is in place State-wide, value-added assessments should be made for all schools, districts, and the entire State. For high school students, a more complex measure must be developed since the children do not take annual tests but rather Regents exams. In these cases, to mirror the “value-added” approach used for the earlier grades, some comparison must be made of Regents scores of the high school’s entering class to their grade 8 exams.

Individual schools can be held accountable only for their value added, since they have little control over what children have learned before they enter the school. This is especially true for middle and high schools, but germane even for elementary schools where what children know before entering school can have important implications for their performance in school. In addition, since students move between schools, such school-based systems will only partially assess the performance of the schools and the district. However, an entire district can be held accountable not only for the value added by its schools, but for whether the entire system has brought its students – most of whom remain in the district – to an adequate level of performance.

- ***Districts should also be held accountable for the absolute level of achievement of their students on standardized tests rather than simply for their rate of improvement.***

The accountability system for K-12 districts should reflect the average difference from Level 3 or from Regents standards as an absolute measure of performance.

Most testing research also endorses the notion that children should be tested using various methodologies, and at various times, to avoid over-weighting a single test and a single style of learning.

- ***Statistically valid methods of assessing student achievement other than through standardized tests should also be incorporated into the accountability system.*** Districts and the State should be encouraged to develop such alternative assessment mechanisms – such as structured portfolio assessments – for inclusion in the accountability system.

Quantitative evaluations: other tools

Other measures of the performance of schools and districts offer additional information about the quality of the educational services they deliver.

- ***Accountability systems should incorporate other quantifiable measures of performance.*** Data should be included in the State accountability system that reflects such critical elements of the educational process as the safety of the students, the level of parent engagement and satisfaction, attendance, and the quality of the school facilities. State measures for these criteria should be identified and reporting required by all schools and districts in the State. Schools and districts should be held accountable for year-over-year improvements and for progress towards State-established goals.

Qualitative evaluations

Finally, no system of assessment based simply on quantitative data can adequately measure the performance of complex institutions such as schools. Indeed, institutions such as hospitals are required to undergo regular assessments by outside experts that include extensive site visitation and interviews, as well as data review.

- ***Qualitative assessment programs should be implemented as part of the State-wide accountability system.*** The State should include a qualitative assessment program in its accountability system. On-site visits need not take place annually, but on a rotating three-to five-year cycle. They should be performed objectively by officials from the State Education Department, teachers and leaders from other districts in the State, and pedagogical experts from the leading education schools. No school should be subject to significant sanctions or removed from a sanctioned status without such a site review verifying the challenges faced or the progress made. To keep the number of schools visited each year to a manageable number, the assessments could be focused on schools at risk of entering or leaving sanctioned status, with a randomized sample of all other

schools used. The State should also consider using an external entity to organize the assessment program to ensure its credibility.

Efficiency accountability recommendations

The interests of the taxpayers in the State need to be protected as well as the interests of its children. The funds used for instruction must not only be used effectively – the mandate of the Court – but also efficiently. Nothing has proved more vexing in the field of education research than finding a clear relationship between inputs of resources and outputs of student learning, and the absence of any consensus on the nature of such linkages (often called “the production function” in the research) should lead to modest expectations for accountability measurements in this vein. That does not mean there should not be such efforts, if only experimentally at first.

- ***The State accountability system should include measures of the relative efficiency of how resources are used by schools and districts.*** Since no absolute standards for efficiency exist in the field, accountability measures for efficiency should be relative in nature. At least two tools should be developed by the State to ensure that some focus on efficiency remains in the instructional delivery system. A ratio of inflation-adjusted per-pupil spending – adjusted for various demographic and need categories – to “value-added” gains should be reported by schools. This ratio can then be compared to how other similarly situated schools in the district performed (so that regional cost differences are muted) and to other schools across the State (with an adjustment for regional cost variation). In addition, even in the event that these comparisons prove problematic, schools should also report changes (adjusted for inflation) in the ratio over time so that the impact of the new funds can be tracked. Schools demonstrating substantial changes in resources with no concomitant changes in student performance should be subject to intervention, at first by the district and eventually by the State.

Districts also should be held accountable for how they use the new resources as well as their current funding. The percentage of total funding used by the district for classroom instruction, school administration, student support services, and district administration should be reported on a consistent basis. Districts can then be compared to each other and with themselves over time to ensure pressure remains to use resources efficiently as possible in the delivery of direct services to students.

Other measures of efficient resources utilization include the use of the physical plant. Schools and districts should be required to report space available per student, which can then be adjusted for factors such as special education (which may require more space per student). Then schools and districts can be compared to each other and best practices in space utilization can be shared.

Consequences

Accountability systems work best when there are clearly defined and predictable consequences that are applied fairly but not mechanically. Furthermore, consequences should be

positive as well as negative, available for leadership as well as staff, and distributed in a timely manner so that the cause-and-effect relationship will be clear. Most critically in the case of schools, sanctions must be carefully targeted so as not to punish the children for the failures of the adults leading their school. The current mélange of accountability systems in the State meets few of these tests.

- ***Clear and reasonable State-wide sanctions for schools should be developed and implemented consistently.*** Sanctions of schools for failing to meet outcome accountability measures should be clear and graduating. They should be consistent with the federally-mandated NCLB actions, if those steps can be implemented by the State sensibly and manageably (it is very unclear whether a system that requires sanctioning of as many as half the schools in the State is meaningful or effective). Sanctions for failing to meet efficiency accountability standards should include additional oversight by districts (and in the case of districts, by the State itself).

In addition, the State accountability system should include sanctions for districts failing to meet progress goals. These should parallel those for schools, beginning with additional oversight of failing districts by the State and leading to eventual State take-over.

- ***Clear and reasonable rewards for good performance should be developed and implemented consistently.*** Concurrent with this more aggressive use of the sanctions, the State should develop rewards for schools and districts that perform well. Districts should be encouraged to negotiate monetary school-wide performance bonuses for outstanding work, and non-cash incentives such as State-wide professional recognition and “master principal” designation should also be established to reward leaders whose schools add value for their students. The current NYC administrators’ bonus model can be used as a starting point if expanded to include all teachers in the school, but it needs to be simplified, clarified, and made more timely.

School districts with outstanding performance should also be eligible for reward, and in this case the State should consider making significant cash bonuses to superintendents with very high achievement levels (since the total cost to the State would be small).