

On New York's Assessment Policy: A Perspective from the Field

Margaret Dwyer¹

White Plains Public Schools & Teachers College, Columbia University

ABSTRACT

Starting in the early 1990's, New York State began to establish learning standards that defined what students at various grade levels should know and be able to do in most curricular areas and a series of state assessments designed to measure student progress towards the performances demanded by the standards. Experience with these assessments has provided for practitioners (that is, teachers, administrators, and other professional-level public school workers) the basis and motivation for a critical field perspective on state assessment policy. By a critical field perspective, I mean a perspective taken by practitioners (as opposed to specialists in testing) that is grounded in analysis that is committed as much to equity as it is to excellence in public education. The purpose of this article is to outline such a position on the state assessments. I begin by describing the assessments as they are commonly experienced by practitioners, identifying the political context in which our state assessment policy has developed, and questioning the fundamental nature of the policy itself. I then suggest actions for practitioners to take to challenge problematic aspects of the state policy and to promote policy that enhances both equity and excellence.

INTRODUCTION

Starting in the early 1990s, New York State began to establish learning standards that defined what students at various grade levels should know and be able to do in most curricular areas. At the same time, a series of state assessments designed to measure student progress towards the performances demanded by the standards was phased in. Since then, the New York State Standards and Assessments have become a common denominator for all public schools in New York State, whether the schools are upstate or down; urban, rural, or suburban; well-financed or under-resourced. Experience with these assessments has provided for practitioners (that is, teachers, administrators, and other professional-level public school workers) the basis and motivation for a critical field perspective on state assessment policy. By a critical field perspective, I mean a perspective taken by practitioners (as opposed to specialists in testing) that is grounded in analysis that is committed as much to equity as it is to excellence in public education. The purpose of this article is to outline such a position on the state assessments. I begin by describing the assessments as they are commonly experienced by practitioners,

¹ Margaret Dwyer has worked to promote the interests of language minority learners for over twenty years as a classroom teacher, a schools administrator, and as a teacher educator. Correspondence concerning this article should be sent to Margaret Dwyer, Director, Newcomer Center, 350 Main Street, White Plains, NY 10601. E-mail: margaretdwyer@wpcsd.k12.ny.us.

identifying the political context in which our state assessment policy has developed, and questioning the fundamental nature of the policy itself. I then suggest actions for practitioners to take to challenge problematic aspects of the state policy and to promote policy that enhances both equity and excellence.

CONSIDERING OUR COMMON EXPERIENCE

Our common experience with state assessment policy is complex. It is undeniable that a good program of assessment is an essential and powerful tool for improving teaching and learning, and is thus highly desirable. On a classroom level, for example, we know that we can neither begin to teach at an appropriate level, nor monitor the effectiveness of daily lessons, nor be sure that students have truly learned what was taught, without careful assessment. A solid statewide assessment program that is connected to a shared set of worthy curriculum standards could provide us with a powerful tool to improve teaching and learning. It could also be part of creating a more equitable statewide educational program, since it would have the potential to give us information about the degree to which all students are engaged in the same high-quality curriculum across the state.

The movement to create learning standards and assessments for major curricula across the United States has claimed to be an effort to improve teaching and learning for all students (Gandal & Vranek, 2001). In New York it served to eliminate what had been a two-tiered educational system (Yanofsky, 2001). The system that existed in New York before the implementation of uniform state learning standards allowed some children to be engaged in curricula governed by high expectations while the majority were engaged in a curriculum that allowed them to graduate from high school by demonstrating competence on an eighth-grade level. Furthermore, the previous system allowed English Language Learners (ELLs) to be invisible by essentially exempting them from the more rigorous level of assessment while simultaneously failing to provide resources to enable them to attain the higher level (Kilian, 2001). The drafts of the learning standards and the early pilot versions of the assessments promised a new day in education in New York. There were to be high expectations for all learners, and their progress was to be assessed by tests that would engage them in truly demonstrating their learning.

Education policy is, on the one hand, a statement about the kind of society we want, and in this case, the policy is a statement in support of a rigorous public education for all. On the other hand, policies are blunt instruments whose powers are limited by myriad factors (Elmore, 1983). In the case of state assessment policy, we have found ourselves living with the most challenging policy instrument of all – unfunded mandates – which has meant that standards and assessments are imposed with no funding to enable districts to comply with the raised expectations.

Therefore, the reality of the assessments as they are now experienced falls far short of what we had hoped for. The Standards outline high expectations for all students, but we have no targeted funding to minimize disparities in the capacities of local districts to realize these expectations. Moreover, the tests themselves can be a stunning burden. Consider, for example, that the average native English-speaking eighth grader (and it is important to remember that such a student is only twelve years old) takes five separate state assessments, spending the same

number of hours on them as the average candidate for the Bar Exam spends entering the profession of law (15 hours total). ELLs, who are entitled to extended time limits, spend as much as double that. Furthermore, that figure includes only the number of hours that the students spend actually sitting for the tests. The New York State Council of School Superintendents (NYSCOSS) has calculated that the total amount of time spent on state tests is actually much greater. For example, for only one of these tests, the Grade 4 English Language Arts Assessment, the average amount of time involved, including administering of the tests, training to score the test, and actually scoring the test, comes to 156 hours (NYSCOSS, 2001).

In addition to the time that they require, these tests exact other costs. They are a financial burden, as districts scramble to find and pay substitute teachers to cover the classes of the teachers who must devote several days to scoring them. This practice leads to unmeasured instructional effects as children spend days with substitute teachers who usually cannot be as effective in the classroom as the regular classroom teacher. The tests can also place an emotional burden on teachers and children alike. Some grades have already seen the exodus of experienced teachers, who ask to be reassigned to grades in which they can spend more time teaching and less time testing (Goodnough, 2001). This kind of stress takes a toll on everyone's spirit and sense of excitement about school and learning.

Thus, our experience with the state assessment policy is one of promise only partially fulfilled. Let us examine the context in which such an uneven policy has arisen and uncover the policy's true nature.

DESCRIBING THE CONTEXT AND NAMING NAMES

Over the past several years, we have experienced a wave of so-called "standards-based reform" across the nation. (I say "so-called" because, as I will explain later, a great deal of the reform that has been called standards-based has not been what I would consider to be truly standards-based at all.) This movement has taken place in a political climate that has been hostile to public education (Berliner, 1999; Berliner & Biddle, 1993). The increasingly conservative political mood in our country is one that glorifies a free market and longs for an idealized traditional past without acknowledging the class and race implications of such positions. Michael Apple has described what has happened in our society as the hegemony of the right in the public discourse. He defines hegemony as "...a process in which dominant groups in society come together to form a bloc and sustain leadership over subordinate groups" (1996, p. 14). A key to understanding hegemony is realizing that coercion is not a necessary element; in fact, hegemony exists when subordinate groups consent to the leadership taken by the dominant groups, usually because the dominant group has persuaded them of the rightness of their ideology. In the public discourse about education today, the right has defined our educational debate and taken over the language of education. As an example, consider the word *standards* itself. It is always used with words like *high*, or *rigorous*, or *world-class*, with the result that to question the standards movement is to be for bad education. Of course, this is not a choice at all, and this is exactly what hegemony looks like.

A first step in getting the kinds of assessment policy we need is challenging the hegemony of the right on subjects related to public education. In so doing, I am not suggesting that we advocate a return either to a laissez-faire educational environment in which there were

very few standards of any type, or to the previous two-tiered system described above. Such systems would find few supporters today, and rightly so. Let us instead claim our place in the dialogue and insist on *true* standards-based reform, which is not, I believe, what reform in New York has been about.

In New York, and in many other states as well, educational policy has been focused on higher academic achievement driven by high-stakes assessments. What we have in New York is not so much genuine standards-based reform, but what Scott Thompson has called its evil twin, “test-based reform” (2001). It is essential for us to be clear about this distinction and to name our policy for what it truly is.

Test-based reform is a system that claims accountability as its goal. In test-based reform, what matters most is improving test scores, which are used to evaluate student learning and school quality. The high-stakes tests result in a narrowing of the curriculum to only the information and skills likely to be tested, and a reduction of instruction to test prepping. Lost in test-based reform are precisely the kinds of rich teaching and learning experiences that school reform is purported to promote. As the Chief School Administrators of Southern Westchester [County] have noted, “...in so many schools, the joy of learning is evaporating in this atmosphere of frenzied preparation of children for these new exams” (2000, p. 6).

One sure sign that the type of policy we have in New York is test-based rather than truly standards-based is the fact that the most outspoken critics of the policy are some of the wealthiest school districts in the state, where test scores are very high. Although they have many objections to the policy, one of the most important that they cite is exactly what Thompson finds so unacceptable about test-based reform. Quite simply, they are able to engage in the type of test preparation that produces good scores, but they do not see the value in it (Empire State Supervisors and Administrators Association, 2001). So outraged about the flawed policy was one wealthy district, Scarsdale, that it encouraged parents to boycott one of the assessments in the spring of 2001. Scarsdale Superintendent Michael V. McGill communicated his position on the state assessments by writing to parents, “Excesses of the standards movement have promoted lock-step education. They’ve diverted attention from important local goals, highlighted simplistic and sometimes inappropriate tests, needlessly promoted similarity in curriculum and teaching. To the extent they’ve caused education to regress to a state average, they’ve undermined excellence” (Zernike, 2001).

In stark contrast to narrow test-based reform is true standards-based reform with equity and excellence as its goals. If test-based reform is about accountability, then true standards-based reform is about quality for all. In this type of reform, we have not only curriculum and performance standards, as we now have in New York, but also what the National Council on Education Standards and Testing (NCEST) has called “opportunity-to-learn standards” (Darling-Hammond, 1997). Such standards refer to steps that are taken to ensure that every child is provided with at least the minimum conditions that would be needed to reach the expected performance standard – things like access to a certified teacher, adequate textbooks, and sufficient time and exposure. Performance standards in the absence of opportunity-to-learn standards can actually exacerbate inequality. An immediate, obvious example is the inequity of a system that mandates a high level of performance in the sciences, and assumes laboratory and technology skills and a broad knowledge base, but that does not also mandate and provide for equal access to laboratories, computers, and certified, qualified teachers.

Let us consider what opportunity-to-learn standards might mean for ELLs in New York State. As a start, providing ELLs with an opportunity to learn that is equitable would mean

ensuring that all ELLs (and all children) are taught in classes of reasonable size by certified teachers with access to adequate instructional resources regardless of the relative wealth of the community in which they are schooled. It could also mean that instructional programming decisions would be based on our best findings in educational research, and that, as a result, various parts of an ELL's school day would be devoted to native language development and enhancement, English language development and enhancement, and rigorous content area study in whichever language was appropriate. Opportunity-to-learn standards would require *equitable* as opposed to *equal* access to learning opportunities. Thus, they would also almost undoubtedly mean high-quality instructional opportunities beyond the normal school day and over the summer in order to provide ELLs with time and exposure sufficient for a non-native speaker of English to demonstrate mastery of content area subjects and English after a reasonable length of time.

As we explore the context in which our state assessment policy has evolved, we need to be clear that what has been called standards-based reform is in many states really test-based reform. Further, a critical perspective on state assessment policy includes the awareness that student performance standards enforced through high-stake assessments in the absence of opportunity-to-learn standards are unjust, and that this can result in penalizing children for not knowing what they have not been given the opportunity to learn.

A CALL TO ACTION

Even within the existing policy environment there is the possibility for creating a model of education that is just and educationally sound. Its seeds are to be found in the possibilities of positive change that are to be found in true standards-based reform, and its promise can be realized by practitioner/activists acting in coalition.

Kohn (2001) reworks a famous idea to remind us that all that is necessary for the triumph of damaging educational policies is that good educators keep silent. Perhaps the single most important action that concerned educators can take is to take on activist roles within our professional organizations. Such organizations would surely include the obvious advocates for the interests of ELLs: New York State Teachers of English to Speakers of Other Languages (NYSTESOL), and New York State Association of Bilingual Educators (NYSABE). However, most practitioners also belong to professional associations that wield significant political power in the state, and we limit our effectiveness when we do not engage associations such as New York State United Teachers (NYSUT), New York State Association of Supervisors and Administrators (NYSASA), and NYSCOSS in our lobbying efforts. Just as we need to refuse to allow ESOL to be a marginalized program in our schools, we need to refuse to marginalize ourselves politically by missing opportunities to work with other powerful allies who have as part of their mission the representation of the interests of all children.

As part of this broad and powerful coalition, we need to lobby on a statewide level for policy changes both profound and mundane. Clearly, the most basic and essential is targeted funding that would create more equal opportunity to learn. Such funding would enable districts to make the playing field more equal across the state by creating more equal access to qualified teachers, good textbooks and equipment, and extended programs after school and in the summer.

A second profound change needs to take the form of humane assessment guidelines that, for example, do not overburden any one grade, and that allow children sufficient time and exposure to learn the curriculum. This would eliminate the clustering of tests at the eighth grade level mentioned earlier, and it would also acknowledge that New York's schools have thousands of newly arrived ELLs who have not yet had access to New York's curricula.

If the tests are to be a useful tool in improving teaching and learning, then we need data that are timely and clear. Current procedures allow for a lag of several months between test administration and results notification. If data are not timely, then assessment is not useful in informing and improving instruction, and they become useful primarily as a tool for criticizing public education. Districts should demand and receive clear, useful test results shortly after test administration. Such practice would allow administrators and teachers to work together to interpret the data, to uncover strengths and weaknesses in the instructional program, and to plan for interventions in the instruction of children who are not making adequate progress. Conversations about worthy, timely data should be at the heart of efforts to improve instruction. The failure of the state to live up to its policy promise of an assessment program that will inform and improve instruction is unpardonable.

We also need transparent test development procedures that inspire confidence. On the macro level, we deserve to know how the tests were designed, and we should expect evidence of validity and reliability. On the micro level, if we are to have reasonable interrater reliability, we need unambiguous scoring guides with clearly identified criteria that minimize subjectivity, and adequate lead time to allow raters to interact with the guides themselves and with the exemplars. (This actually is a good problem to have. It is indicative of the fact that we now have some performance-based tasks on our tests, not just cloze or multiple choice. However, these tasks require a higher level of sophistication in scoring, and we need to demand of the State Education Department that they contribute to building this capacity.)

Finally, we need to demand more accurate reporting methods that emphasize growth in individual schools rather than competition among schools and neighborhoods that may be very different. Such reports also need to differentiate among scores for children who have and have not had access to the curriculum in New York. If, for example, we are going to require that newly arrived ELLs take the state assessments whether or not they have had access to any instruction in the curriculum, then their scores should be indicated as base-line scores, not, as they currently are, as "failing to meet the standards." In the absence of evidence of reliability of the tests themselves, we need scores that are reported along a band rather than as discrete numbers. Concerns about the reliability of the tests have been repeatedly raised from the field (e.g., Chief School Administrators of Southern Westchester, 2000). Although the State Education Department has yet to provide evidence of reliability, they allow scores to be publicly reported as if they were accurate reflections of an absolute level of achievement.

Not everything that needs to be challenged is in Albany. There is a lot we can do in our own districts to be smarter and more strategic in the ways that we interact with the state's policies. On the level of the school district, there are actions we can and should take to limit the power of the assessments, thus mitigating the high stakes that they command. We can, for example, simply refuse to base high stakes decisions such as promotion or retention to any single assessment. The assessments can certainly have a place in our decision-making process, but they should be used only in combination with a broad menu of criteria that collectively paints an accurate picture of a child's learning.

We also can seize professional development opportunities by providing teachers with clear data, and help in interpreting and responding to them. For example, data that reveal to a teacher that her students consistently fail to organize their thoughts effectively in a pre-writing graphic organizer can point her practice in ways that will enhance student performance. As school leaders, we can also analyze test data on a grade-level or school-wide basis and create opportunities for teachers to learn more effective practices in targeted areas.

It is also important to confront the local press about sensationalist, shallow, and downright inaccurate representations of data. Student test results are published as program evaluations by local papers that do not ask the hard questions about the reliability, validity, and true meaning of the test scores. We need to educate journalists about the complex issues connected with the assessment policy and urge them to use their power to educate and advocate for policies that are more beneficial to all learners. The best forum for this education of journalists is not at the school district level, however. In order for this undertaking to be taken seriously by the press, it should be based in our most prestigious schools of education, where expertise in assessment is perceived to exist, and where personal motivation cannot be attributed. If the level of public discourse about assessment policy is to improve, we need graduate students, researchers, and professors of education and educational assessment to take the lead in organizing press conferences and publishing brief, accessible monographs for journalists on issues related to assessment policy.

CONCLUSION

Practitioners grappling with current New York State assessment policy find ourselves living with an unfulfilled promise. While we have made progress towards eliminating in theory the two-tiered educational system of the past, in practice we find ourselves living with unfunded mandates that promote test-based reform rather than true standards-based reform. As practitioner-activists, we are ethically bound to confront the policy critically and to struggle for educational reform that truly does keep the promise of excellence and equity for all.

REFERENCES

- Apple, M.W. (1996). *Cultural politics and education*. New York: Teachers College Press.
- Berliner, D.C. (1999). Julius and Rosa Sachs Lecture, Teachers College Columbia University, Spring 1999.
- Berliner, D.C., & Biddle, B.J. (1995). *The manufactured crisis: Myth, fraud, and the attack on America's public schools*. Reading, MA: Addison-Wesley.
- Chief School Administrators of Southern Westchester. (2000). *In support of higher standards; Concerns about implementation*. (Position paper for the NYSED, March.)
- Darling-Hammond, L. (1997). *The right to learn*. San Francisco: Jossey-Bass.
- Elmore, R. (1983). Complexity and control: What legislators and administrators can do about implementing public policy. In L. Shulman & G. Sykes (Eds.). *Handbook of teaching policy*. White Plains, NY: Longman.

- Empire State Supervisors and Administrators Association. (2001). Scarsdale steps up to the plate. *The ESSAA advocate*, II/1, Spring, 1.
- Gandal, M. & Vranek, J. (2001). Standards: Here today, here tomorrow. *Educational leadership* 59/1, 6-13.
- Goodnough, A. (2001, June 14). Strain for fourth-grade tests drives off veteran teachers. *New York Times*, A1&B4.
- Kilian, L. (2001, September). Personal communication.
- Kohn, A. (2001). Fighting the tests: A practical guide to rescuing our schools. *Phi Delta Kappan*, 349-357.
- New York State Council of School Superintendents (NYSCOSS). (2001). *Assessing the assessments*. (Report of the Curriculum and Instruction Committee, June.)
- Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, 358-362.
- Yanofsky, S. (September, 2001). Personal communication.
- Zernike, K. (2001, April 13). In high-scoring Scarsdale, a revolt against state tests. *New York Times*, A1 & B4.