

The Impact of England's National Curriculum and Assessment System on Classroom Practice: Potential Lessons for American Reformers

DAVID L. SILVERNAIL

Today many American educational reformers and policymakers are beginning to embrace a new reform strategy for improving our schools, a strategy that emphasizes results rather than inputs or the redistribution of power. Key components of this strategy include the creation of "world-class" standards, curriculum frameworks, and "voluntary" national tests. Reforms similar to these are already in place in England. The 1988 British Education Reform Act provided for the creation of a national curriculum and national assessment system. This article describes the British experience with these reforms and identifies potential lessons for American reformers.

WITH the passage in 1994 of Goals 2000: Educate America Act, American education faces yet another wave of educational reforms. *A Nation at Risk*

AUTHOR'S NOTE: This research was supported in part by the National Center for Restructuring Education, Schools, and Teaching (NCREST) at Teachers College, Columbia University. The views expressed in this article are those of the author and do not necessarily represent the views of NCREST. An earlier version of this paper was presented at the 1994 annual meeting of the American Educational Research Association in New Orleans, Louisiana, and a more detailed report of this study is available from NCREST.

EDUCATIONAL POLICY, Vol. 10 No. 1, March 1996 46-62
©1996 Corwin Press, Inc.

(National Commission on Excellence in Education, 1983) set into motion what is commonly referred to as the first wave of reforms of the 1980s. State legislatures and departments of education in all 50 states passed legislation and regulations that reflected a top-down approach to reform. By one account, over 700 pieces of legislation affecting education were passed between 1984 and 1986 (Futrell, 1989), legislation that increased graduation requirements, extended the school day and year, required minimum competency testing, and toughened teacher-certification requirements.

Publication of *A Nation Prepared: Teachers for the Twenty-First Century* (Carnegie Forum on Education and the Economy, 1986) was seen by many as ushering in the second wave of education reforms in the past decade. These reforms were premised on a bottom-up approach to change, a belief that schools must change from the ground up, one building at a time. Beginning in the mid-1980s and continuing up to the present, educators and policymakers alike have been implementing reforms designed to professionalize teaching, to create new leadership roles for teachers, to deregulate schools, and to promote site-based management and school restructuring.

Goals 2000 may very well be viewed as the official beginning of a third wave of reforms, a wave emphasizing results rather than inputs or the redistribution of power. For several years, many have called for "world-class" standards and some kind of national tests, but passage of Goals 2000 marks the formal sanction and increased federal role in creating these standards and assessment. The act codifies eight national goals; establishes the National Education Goals panel and the National Education Standards and Improvement Council; and provides federal funding for the development of voluntary-content, opportunity-to-learn, and student performance standards, as well as new assessment systems. Will these new reforms succeed where earlier ones have failed? Only time will tell, but it is possible to get a glimpse of how these reforms may affect classroom practice by examining England's recent 6-year experience with similar educational reforms.

In 1988, the British Parliament passed the Education Reform Act, a landmark act that created a national curriculum and a national examination system. In many ways, the curriculum and assessments are uniquely British. However, several key components of the British reforms and those recently introduced in the United States are similar enough to suggest that we might find some lessons in the English experience. This article will suggest some of these lessons after first describing briefly the British reforms and exploring the impact they have had on classrooms and teachers.

OVERVIEW OF THE BRITISH NATIONAL CURRICULUM AND ASSESSMENT SYSTEM

The centerpiece of the 1988 British Education Reform Act was a provision for creating a national curriculum and a national assessment program.¹ There are 10 foundation subjects in the national curriculum. These include three core subjects (English, mathematics, and science) and seven *other foundation subjects* (technology, history, geography, music, art, physical education, and, for secondary students, a foreign language). Each subject is broken down into broad categories of knowledge, skills, and understandings, and for each category there is an accompanying *attainment target*, a broad outcome identifying what a student is expected to learn. Each attainment target in turn contains *statements of attainment*, more specific statements (objectives) that describe up to 10 *levels of attainment* (achievement). These 10 levels, which form the basis of assessments, are designed to cover the range of attainments expected of students during their age 5 to 16 compulsory schooling years.

The national assessment program consists of two parts, teacher assessments and standard assessment tasks (SATs). The national assessments take place at or near the end of four key stages, stages that correspond to four ages (7, 11, 14, and 16 years old). Teachers first conduct a teacher assessment of each child's accomplishments of a particular statement of attainment and then use this assessment information to select the series of SATs they will administer to each child (e.g., test packets covering Levels 2 to 5, Levels 3 to 7, etc.). Scores on SATs are used to set (establish) a child's level of attainment.

Once completed, SAT scores are combined with the teacher assessments and compiled into a report for the individual child's parents. A second report that describes individual school results is submitted to the national Department of Education. The parent reports are intended to inform parents of their child's progress, whereas the government report is intended to become part of what are called *league tables*. League tables report the schoolwide assessment results for each school in each LEA and a school ranking based on the school's assessment performance.

IMPACT OF THE NATIONAL CURRICULUM AND ASSESSMENTS

What has been the impact of the British national curriculum and assessments on classroom practice? This question was put to over 50 teachers, headteachers, educational officials, and policy researchers who were interviewed in England over a 2-month period in mid-1993. The evidence from these extensive interviews is clear. In just a few short years, the implementation of the national curriculum and assessment system has had a significant

impact on British education. For an overwhelming majority of teachers, these reforms have fundamentally changed the way they teach, what they teach, and how they assess their students.

British researchers have found that a majority of England's teachers support the concept of a national curriculum (see Cox, Evans, & Sanders, 1992; Silcock, 1990). As one teacher remarked, "The introduction of a national curriculum has resulted in a general tightening up and systematic approach, which was long overdue."² However, as one begins to probe this teacher support, a very different picture surfaces. Many of the teachers, even those who support the concept of a national curriculum and assessments, believe that the specific national curriculum and assessment program, as it has been implemented, is detrimental to the education of British children. Why? As one educator so aptly put it, "The devil is in the procedure."

Where is the devil? It is in many places. It is in the process itself that has been used in implementing the national curriculum and assessments. It appears in what gets taught, how it gets taught, and who gets taught. And it appears in what gets assessed.

First, the problems of implementation. The reform act was passed by Parliament in 1988 and implementation of the national curriculum and assessments began immediately. The national curriculum in the core subjects was put in place by autumn 1989 and other subjects have followed in close succession. The Task Group on Assessment and Testing (TGAT), the group charged with developing the national assessment system, was given only 5 months to develop the assessment system, and after field-testing items in 1990 and 1991, SATs were administered to all students at Key Stage 1 in the summer of 1992. The pace of change has been hectic. Equally important, the sheer number of changes and false starts has created havoc for teachers. A good example is the science curriculum. Originally, there were 17 attainment targets. This was soon reduced to 11, then 5, and most recently to 4. Each time statements of attainment changed, teachers have had to redesign lessons and teacher assessments. The same is true in the other subject areas and with SATs. As one teacher remarked, "It is not the national curriculum, it is the national chameleon!"

The curriculum was intended to provide each child a broad and balanced curriculum. In reality, however, it appears that many children receive neither a broad nor balanced curriculum. England's teachers have made a Herculean effort to implement the national curriculum. Still, the sheer magnitude and detail of the curriculum have forced teachers to narrow what they can teach. This is particularly true at the primary school level. By statute, primary teachers are required to teach each of the 10 foundation subjects. However, many of the primary teachers interviewed indicated that they lacked the full

range of academic background necessary to teach the national curriculum, and many openly admitted they were not teaching each discipline equally well.

Even if adequately prepared, many teachers find themselves unable to cover all 10 subjects. The national curriculum was built piecemeal. For each discipline, there was a working curriculum committee that developed the programs of study, attainment targets, and statements of attainments. There was considerable debate within each committee over what should be included in the curriculum, and in most cases, the committee members compromised on the side of inclusion rather than exclusion. The core knowledge and skills for each subject became enlarged and more expansive. As a consequence, most of the individual discipline curriculums are broad but when added together with the other nine disciplines that are to be taught in primary years, teachers become overwhelmed. One crude estimate was that Key Stage 1 teachers needed to teach and assess approximately 8,000 statements of attainment in any given year.

Given this large number of assessments, teachers have been forced to choose which subjects will get more or less attention in their classroom. A clear indicator of this choosing process can be seen in how teachers divide their instructional time. Campbell and Neill (1992) found that approximately 56% of the teaching time in primary school classrooms was devoted to the core subjects and only 27% to all the other foundation subjects. And it is no accident that a majority of the instructional time was spent on the core subjects, subjects that are tested at Key Stage 1. Teachers felt pressure to target their teaching toward those statements of attainment that are tested with SATs.

Another way in which the national curriculum has narrowed the classroom curriculum is by forcing teachers to dilute what they can teach and how they can teach it. This happens because the scarcity of time precludes studying topics and skills in the detail desired. A typical teacher comment was:

Obviously, children need to revisit the same skills and concepts in a variety of ways before they take them on board. I feel we are no longer able to cover everything we would like to, or feel necessary to do, in as much depth as we feel is justified, because of the pressure to get through an inordinate amount of work in the core and foundation areas.

The national curriculum also dilutes the classroom curriculum by limiting the time available to explore new avenues of learning. Teachers feel they have lost the freedom to run with the "teaching moment." They cannot take time out from the prescribed curriculum to turn an unexpected or serendipitous event into a learning event. A primary teacher remarked, "For example, if a child brings in a bird's nest or a kite, we say 'very nice, dear,' but can no longer spark off ideas brought in by the children."

The teaching curriculum is also being narrowed by the very nature of the content included in the national curriculum. To be more accurate, it is the type of content that is being excluded that narrows the curriculum to which students are being exposed. A primary headteacher in an urban community reports, "The national curriculum is terribly restrictive in history. We have 22 different languages of origin represented in our school. Our school reflects a highly multicultural society, yet almost all of the history [in the national curriculum] is British history." The same narrowness in content can be found in many disciplines and, possibly, most notably in the English curriculum. As one secondary teacher voiced it,

The English curriculum is very prescriptive, prescriptive to the extent, for example, they tell us now that every Key Stage 4 student must study a pre-20th-century novel. And the anthology for Key Stage 3 reflects a desire to return to a time "when things were better," to return to "this rosy period of time." They think children should have the experience of it. They are going to ram it down their throats whether they like it and whether the teachers like it or not. And the anthology [of pieces] is something like 85% masculine and 95% White, English. Preserving the White, European male is the notion!

The prescribed assessment program also contributes to narrowing the curriculum. This occurs in several ways, one of which has already been mentioned. Teaching the core subjects consumes a majority of the available classroom teaching time. These are the very core subjects that are currently tested at Key Stage 1. Thus the assessed curriculum is becoming the taught curriculum.

In addition, teachers report that administering the prescribed assessments, and in particular the SATs, limits the time available for teaching. This is because the SATs must be administered individually or in small groups and each child must be assessed on each statement of attainment. The problem becomes what to do with the rest of the class during administration of SATs. Some schools have been fortunate enough to have sufficient funds to hire a supply (substitute) teacher. Most have not. Many teachers report assigning independent work for the other children in the class, work that is not properly supervised, nor considered useful.

The SATs, as they have evolved over time, have also reintroduced a clear demarcation between the activities of teaching and assessment. Historically, the line between teaching and assessing in English primary schools has been a fuzzy one. Sometimes an activity was used to teach a child and sometimes to assess the child, and sometimes teaching and assessing went on simultaneously. The first SATs were largely activity and performance based, much in keeping with traditional British primary education. Teachers liked the

tasks, but their required administration for every statement of attainment for every child significantly increased teachers' workloads. In addition, administration of the tasks was found to vary considerably from teacher to teacher, and the reliability and generalizability of the individual performance assessments were highly suspect. Consequently, the government revised the SATs and made them more like traditional, objective, paper-and-pencil tests. And because they were to be used as accountability tools, teachers no longer can help the children while they are taking their SATs. They cannot coach the child. As a result, many of the primary teachers interviewed expressed frustration with having to take time away from teaching to stop and administer the SATs; teachers feel they must separate teaching from assessing, a practice contrary to long-standing primary education principles.

The national curriculum and accompanying assessment program have had other adverse effects, particularly in the area of curriculum equity. The national curriculum was intended to be an entitlement curriculum, one that guaranteed that all British children were taught a common curriculum. In large measure, this has occurred. A majority of children are being taught and assessed in at least the core subjects. Upon closer examination, however, one finds that not all children are being given equal access to the curriculum.

One way in which access has become unequal is through the reintroduction of streaming. Streaming is the British equivalent to American ability grouping and tracking. The national curriculum, and in particular the creation of the 10 levels of attainments and SATs, has led in many schools to the differentiating of curriculum for different ability groups. The SATs are grouped differently according to difficulty levels (e.g., 1-3 levels, 2-5 levels, and 3-7 levels) and student ability (e.g., so-called below average, average, and able pupils). Each group of SATs contains different reading passages and different levels of questions. Teachers find themselves preparing children differently for the SATs by teaching toward different levels of attainment for different children. This is particularly true at the secondary level where sets of SATs also cover different content. Consequently, different children are receiving different access to the curriculum and instruction.

Inequalities also have been created for many special needs children and "problem children." Special needs teachers thought an entitlement curriculum would be very beneficial for special needs children. Finally, these children would have equal access and opportunities in curriculum and instruction. In reality, though, the national curriculum has become troublesome for both regular and special needs teachers. The prescribed nature of the curriculum and assessments work against the needs of many special children. One teacher remarked,

In the field of special needs, we are being asked to teach children to jump through totally inappropriate hoops. For example, boys with reading ages between 7 and 8, but aged 14, have to spend their allocation of time with me learning how to answer questions on the Key Stage 3 anthology when this time should be spent learning literacy skills. Because key stages are age specific and programs of study are linked to these stages, some pupils never have time to gain confidence and mastery in subjects.

As a result, some schools are choosing to "disapply" special needs students from the national curriculum and assessments. Disapplication, which is permitted by the 1988 Reform Act, allows headteachers to exempt children under special circumstances from the curriculum and assessment requirements. Special needs students in some regular and special needs schools are being exempted from learning the national curriculum and from being assessed with the SATs. Not all schools are choosing to disapply special needs students, but the problem appears to exist to some degree throughout the country.

In the case of so-called problem children, many also have been denied access to the curriculum. Problem children are defined as children with emotional and behavioral problems. By law, these students may be excluded (expelled) from school, with the individual school having no legal obligation to readmit these children. For some schools, exclusions are being used to exempt students from taking the national assessments, which in turn helps raise the league table scores for these schools. One study puts the number of temporary and permanent exclusions at 66,000 pupils per year, and another found that the number of permanent exclusions increased by 66% between 1988 and 1991 (see Pyke, 1993). The government has passed new legislation to curb this practice, but many teachers and headteachers believe that exclusion will continue to be a problem as long as examination league tables are used to rate schools. Consequently, some children will continue to be excluded from schools and denied access to the national curriculum.

By mid-1993, and under intense pressure from teachers, headteachers, and parents, the government finally conceded that some aspects of the national curriculum and assessments were causing problems and needed attention. The government commissioned a review of the curriculum and assessments, and in December 1993, a final report was issued recommending several key changes (Dearing, 1993). These included reducing the amount of the prescribed national curriculum, increasing the time available for teachers to prescribe their own curriculum, reducing the number of attainment statements that must be assessed, and limiting the national assessment only to the core subjects.

The British government has begun to implement these recommendations, recommendations that appear to have the general but cautionary support of both the education profession and parents. However, time must pass before the full impact of these curriculum and assessment revisions can be seen.

POTENTIAL LESSONS

What are some of the potential lessons to be learned from the British experience? There are many, some more germane than others for American educators and policymakers. It would be easy to dismiss the British experience as not applicable to the United States. Britain has created a very broad, detailed, and prescriptive national curriculum that has been mandated for all schools and classrooms. It has been able to do so for two principal reasons. England has a parliamentary form of government. Once a political party gains a majority in Parliament, it can propose new and radical changes and simply pass them into law. In the United States, on the other hand, we have a constitutional government with an elaborate system of checks and balances. The result is that although the majority party in England can make fairly quick and far-reaching changes, changes in the United States are much slower to gain passage by both the legislative and executive branches of government and most often are compromises between opposing views and strategies of the two branches. Second, the British education system has traditionally been "nationally controlled and locally administered," whereas the American system is "state controlled and locally administered." Thus, whereas English policymakers can change the entire system throughout England by decree, American reformers and policymakers must rely on incremental changes taking place in each state and in over 16,000 school districts across the country.

However, one should not dismiss the British experience too quickly. True, it is inconceivable that the United States would in any foreseeable future mandate one national curriculum and assessment program for all states and schools. Nevertheless, the general public and many teachers do support these reforms. Recent Gallup/Phi Delta Kappa polls reveal that approximately 70% of the public do support requiring schools in their community to use a standardized national curriculum, and over 75% favor using national tests to measure student performance (see Elam, Rose, & Gallup, 1992, 1993). In addition, a recent poll reported that a large majority of teachers support the establishment of national standards for what students should know, and three out of five teachers support a national examination system (Metropolitan Life, 1993).

Currently, several initiatives are under way in the development of these national standards and accountability tests. Supported by either federal funds

or foundation grants, educators in the core subjects named in the national goals (English, mathematics, science, history, geography) as well as educators in other disciplines (civics, social studies, the arts, foreign languages, physical education, health) are developing standards, all scheduled to be published in the next 12 to 18 months. Also, by one recent account, 43 states are currently developing or are ready to implement student content and performance standards (Council of Chief State School Officers, 1994).

On the testing front, the National Assessment Governing Board (NAGB), which is responsible for overseeing "the nation's report card," is attempting to expand the use of the National Assessment of Educational Progress (NAEP) program to provide for state-by-state assessments (and possibly district-by-district assessments) and moving ahead to create achievement levels (basic, proficient, advanced) for gauging student performance. Over 45 states now have mandated student testing programs (Office of Technology Assessment, 1992).

Goals 2000 provides the mechanism and funds for further developing and certifying these various standards and assessment initiatives. Participation is voluntary on the part of states, but states that choose not to participate may soon find themselves marginalized and without access to federal funds. Thus, at the very least, it is highly probable that most states in the near future will have adopted state-mandated curriculum and assessment systems that coincide with federally certified world-class standards and assessments and look remarkably similar from state to state. In effect, we will very soon have a quasi-national curriculum.

A potentially important lesson could be that the creation of this quasi-national curriculum may actually benefit children, particularly if it includes opportunity-to-learn standards. The British national curriculum is designed to be an *entitlement* curriculum, one that provides each child, regardless of ability, family circumstance, and geographic location, instruction in 10 core and foundation subjects. As described earlier, there are some notable exceptions, but for the most part, British children now have an equal opportunity to learn a common body of knowledge and skills.

Adoption of the curriculum frameworks and standards now being developed throughout the United States could serve a similar purpose. It could promote greater entitlement of all American children to a solid, well-rounded core education. Curriculum offerings, materials, textbooks, and standardized tests used by school systems throughout the country are fairly similar. On paper, most school programs look alike. However, it is well documented that the specific curriculum and instruction children receive varies considerably by student ability and achievement levels, and in many cases by gender, ethnicity, and geographic location. All these factors contribute to unequal

access to curriculum content and uneven levels of instruction. Content and performance standards, if appropriately developed, should define what every child should know and be able to do. Also, opportunity-to-learn standards, if appropriately defined and implemented, could ensure that all children have equal access to, and the opportunity of, acquiring this foundational education. A key is in ensuring that these opportunity-to-learn standards are implemented in tandem with the content and performance standards.

What is included in the standards and how they are developed are also extremely important. In England, working groups were established in each of the 10 curriculum foundation areas, and each group worked independently in defining attainment targets, programs of study, and statements of attainment. Conflicts over inclusion of specific content were many times resolved by adding more to the curriculum, not less. No attempt was made to create one unified core curriculum. Rather, each group created its own core, and each core was combined with others to create the national curriculum. The outcome was a curriculum that was too large, too cumbersome, and too prescriptive. Instead of creating a broad and balanced curriculum as envisioned in the 1988 Reform Act, the curriculum that was taught became very narrow.

The same outcome could occur here in the United States. Many discipline organizations and statewide groups are creating national standards and curriculum frameworks. The creation of the standards and frameworks has great potential, but how they are developed in practice needs careful attention. Preliminary working materials produced by several of these groups suggest that each discipline will have numerous standards, and at present, there is no effort under way to coordinate the development of these standards and curriculum frameworks across the various disciplines. They are being developed in a piecemeal, unorganized fashion. No attempt is being made to determine the compatibility of standards across disciplines, nor to determine if the standards when put together will be too broad and unmanageable. Consequently, in the very near future, American teachers may face problems similar to those faced by their English counterparts. They may have to narrow their own classroom curriculum as they attempt to achieve all the new standards.

This uncoordinated proliferation of standards and frameworks does not have to occur. The British government did not anticipate this problem, and, hence, only now after much damage has been done are steps being taken to streamline the curriculum into something that is manageable and realistic. We could learn from this experience. Coordinating committees could be created now and given the task of making sure that the various discipline standards are compatible, reasonable, and achievable within the constraints

of the classroom. This coordinating role could be a responsibility of the National Education Standards and Improvement Council (NESIC) as they review and certify standards, or each state may wish to establish their own council. Whatever the case, practicing classroom teachers should have an important voice in these deliberations. If not, it appears we may well follow in Britain's footsteps and create standards and curriculum frameworks that are unachievable and deleterious to good classroom practices.

England's recent experience with national assessments may also be instructional in several ways. The national assessment system is in considerable disarray. A central cause for many of the problems may be attributed to a shift in purpose of the assessments. The TGAT recommended that the national assessments be used principally for formative purposes and that they resemble normal classroom work. The assessments were to be designed to assess pupil progress and assist teachers in identifying subsequent appropriate learning activities. However, the Conservative government had a different purpose in mind. They wanted to make schools and teachers more accountable, and they wanted to use the assessment results to raise standards. They wanted to use assessment to drive reforms (see Black, 1994). The result was the creation of high-stakes assessments for teachers. Teachers very quickly realized the SATs were not to be viewed as learning and teaching tools but as assessments of their performance.

The underlying premise of England's assessment-driven reform strategy was that high-stakes assessments would motivate teachers and headteachers to improve instruction and student achievement. The same logic undergirds many of the calls in this country for new forms of national tests and national assessment systems. Several studies conducted in the past few years have documented how high-stakes traditional tests result in narrowing the curriculum, encourage teaching to the test, and in some cases, lead to unethical behavior on the part of teachers and schools (see Darling-Hammond & Wise, 1985; Shepard, 1991; Smith & Rottenberg, 1991). Many advocates of national tests and assessment systems appear to believe that these problems can be avoided if the tests are more "authentic." However, one of the most important lessons to be gleaned from the British experience is that this is not the case. High-stakes tests, whatever their form, are still high stakes. They are very vulnerable to what Madaus (1994) called *peiragenics*, test-induced illnesses that create unintended negative consequences.

Is it possible to create assessment-driven reform programs without *peiragenics*? Possibly. One model may be found in the Scottish education system. Scotland has a national curriculum and assessment system. However, in contrast to England, the sole purposes of the assessments are formative and diagnostic—to aid teachers in the teaching and learning process. By law,

the assessments may not be used to provide comparisons between schools. The results are that the assessments have low stakes and greater potential for actually improving learning and student performance.

By shifting the purpose of the SATs to one of high-stakes accountability, the British government created several concomitant problems. One is the problem of attempting to use a single assessment for more than one purpose; in this case attempting to use SATs for both formative and summative purposes. British teachers report that the administration of the SATs significantly reduced the amount of time available for teaching and learning. This was the case in part because of the sheer number of assessments they had to perform. In addition, by attempting to use the SATs for summative purposes, this altered their potential use as formative assessment tools.

Assessments designed to promote and enhance teaching and learning, if properly used, are an integral part of the classroom curriculum and instruction. They are formative and diagnostic and help both the student and teacher understand what the student knows and is able to do now, and *how well* the student knows and is able to do something. They may or may not be formal assessments and often involve a dialogue between the teacher and student. They are developmental. They oftentimes may be modified during the assessment to capitalize on what it appears the student knows and is having problems understanding. These types of assessments provide teachers opportunities to coach students in their learning. As such, assessment becomes another teaching strategy in the teacher's repertoire of instructional practices, and the line between teaching and assessment becomes less defined.

Assessment in British classrooms, particularly at the primary level, have traditionally had these formative assessment features. However, England's teachers found that they had to change their assessment practices as the emphasis of the SATs shifted to summative accountability. Accountability measures are designed to provide a snapshot of a student's achievements—to provide summative statements about what a student knows and does not know and can and cannot do at any one point in time. To be valid, the assessment tasks, administration, and scoring must be formalized and standardized. In addition, to be valid measures of individual student behavior and to permit comparisons across groups, accountability tests must be solo exercises. Teachers must first teach, then stop and administer assessments. They cannot coach the learner or adapt the assessment midstream so that it becomes more useful in diagnosing learner needs. British teachers report having to make these types of changes when administering the SATs. To provide objective results, teachers report having to remove themselves from the assessment process. They could no longer use the administration of the SATs as an instructional strategy. As SATs became increasingly important as an account-

ability tool, they created an artificial division between teaching and assessment and between learning and assessment.

The shift in purpose also had an impact on the nature of the SATs. The SATs were originally designed to be performance based and to mirror normal classroom instructional practices. What the British government has learned is that when you attempt to use performance-based assessment on a large-scale basis, to assess all students in many content areas, it becomes unworkable. Any performance-based assessment requires more time—more time to administer, more time to score, and more time to interpret. Multiply this by multiple performance tasks, multiple content areas, and multiple students in a classroom, and the workload is enormous. And, like many researchers in this country, British reformers found the validity and generalizability of the performance-based SATs questionable. The British government's solution to these problems has been to make the SATs more like traditional paper-and-pencil assessments.

At present, it appears that many of the assessment reforms that are being advocated and developed in this country are strikingly similar to those found in England. Approximately 40 states attempt to use their statewide testing programs for both accountability and instructional purposes. Nearly all states use or are planning to use writing samples to assess children, and over 20 are developing and implementing alternative forms of assessment, including individual performance assessments, demonstrations, and portfolios (Barton & Coley, 1994). In addition, 4 out of 5 states with these statewide testing programs test *every* child at two or three different grade levels.

American reformers would do well to learn from the British experience. Steps could be taken now to avoid some of the pitfalls of the British national assessment program. First, we should heed the advice of the National Council on Education Standards and Testing (1992). Instead of developing single assessment tools and attempting to use them for multiple purposes, we should develop assessment *systems*, systems that include multiple measures of performance that can, at least ostensibly, serve multiple purposes. These systems should include new forms of assessments, but ones that are valid, reliable, and fair. This will not be an easy task, but one in which substantial progress is being made by researchers in this country.³

Second, reformers should give considerable thought to using some type of sampling procedures in their assessment systems. Two possibilities are either item sampling where different students get different samples of test items, or what Popham (1993) called a genuine matrix sampling procedure, which includes both item and student sampling. By using a sampling procedure, workloads and costs are cut substantially. Equally important, it may reduce the ill effects of potentially high-stakes accountability assess-

ments. Teachers feel less threatened by how the results may be used because with the exception of schools with a single classroom per grade level, schools and not individual teachers become accountable for student performance. A sample of test items and a sample of students can be used to profile how a school is performing. Furthermore, these school profiles may be useful for some instructional as well as accountability purposes. Individual student needs cannot be identified, but schoolwide needs may be identified and remediated. One project that is finding some success in this area is Boston College's Urban Development Assessment Consortium (UDAC). Working with teachers and administrators in a number of urban schools, Boston College researchers have created a school-level alternative assessment system—one that includes NAEP items, extended-response items, and performance assessments. Although the program tests all students in a grade level, item sampling does reduce the amount of time needed to test multiple subjects and multiple skill and knowledge areas while producing schoolwide assessment that teachers have found useful for changing their programs and achieving improved student performance.⁴

A final lesson from the British experience is, in fact, a very old one. Haste does indeed make waste. In its rush to implement a national curriculum and assessment program, England created an unmanageable curriculum and an ineffective assessment system. The curriculum became too large to teach, teachers were unprepared to teach all the subjects, and teachers had to narrow the curriculum and change their instruction to provide "coverage" of the curriculum. The national assessments began to drive instruction, reduce teaching time, increase workload, and provide questionable results. Up to now, the American reforms have been developing slowly. The first set of standards, the National Council of Teachers of Mathematics (NCTM) standards, were released in 1989 and only now are gaining widespread use throughout the country. However, this is very likely to change in the near future. Some funds are already available for states to begin the process of creating standards and assessments, with additional funds likely to be available in the coming years. If we learn nothing else from the British experience, we should recognize that rushing the development of any new set of standards or assessments will in all likelihood create undesirable as well as desirable outcomes, and unintended as well as intended results. Thus we should move slowly and cautiously with our reforms. We should study the effects, both intended and unintended, of introducing new standards and assessments into our classrooms. Fortunately, there is an opportunity for this to occur now. Systematic and thorough evaluations of the impact and outcomes of the NCTM standards on classroom instruction and curriculum, teachers, and students could provide valuable

guideposts to states and organizations as they begin implementing the federal reform agenda.

In summary, American education is in the midst of many reforms, some very similar to those found in England. Yet, before we follow in the footsteps of our neighbors from across the Atlantic, it is important for us to learn from the British experience with these reforms. As Madaus and Kellaghan (1991) warned us:

If one decides to select a feature of another educational system in search for solutions to one's own educational problems, one should ensure in the first place that one understands the functioning of that feature in the foreign system (including any problems that have been identified in its operation), and, in the second place, that the proposed transplant is likely to be compatible with the host system. (p. 23)

Some early lessons have been described here, and it is hoped that these lessons may be useful in identifying what paths to avoid and which ones to follow as we attempt to improve our schools in the coming years.

NOTES

1. The 1988 Education Reform Act applies to both England and Wales. Scotland and Northern Ireland, the remaining members of the United Kingdom, have also instituted educational reforms in recent years, some of which are similar to those in the 1988 act. This article discusses the impact of the act in England only.

2. Unless noted otherwise, all quoted statements in the text were made during taped interviews with the author.

3. See, for example, the work of Baker, Linn, and their colleagues at the Center of Research and Evaluation, Standard, and Student Testing (CRESST) at the Graduate School of Education, University of California at Los Angeles.

4. For more information about the Urban Development Assessment Consortium, contact John Cawthorne at the Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

REFERENCES

- Barton, P., & Coley, R. (1994). *Testing in America's schools*. Princeton, NJ: Educational Testing Service.
- Black, P. (1994). Performance assessment and accountability: The experience in England and Wales. *Educational Evaluation and Policy Analysis*, 16(2), 191-203.
- Campbell, R., & Neill, S. (1992). *Teacher time and curriculum manageability at key stage 1*. London: Assistant Masters and Mistresses Association.
- Carnegie Forum on Education and the Economy. (1986). *A nation prepared: Teachers for the twenty-first century*. New York: Carnegie Corporation.
- Council of Chief State School Officers. (1994). *State baselines for Goals 2000 implementation*. Washington, DC: Author.
- Cox, T., Evans, J., & Sanders, S. (1992). How primary teachers view the national curriculum 1 year on. *Educational Review*, 44(1), 19-29.

- Darling-Hammond, L., & Wise, A. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal*, 85, 315-336.
- Dearing, R. (1993). *The national curriculum and its assessments: Final report*. York, England: National Curriculum Council.
- Elam, S., Rose, L., & Gallup, A. (1992). The 24th annual gallup/Phi Delta Kappa poll of the public's attitude toward public schools. *Phi Delta Kappan*, 74(1), 41-53.
- Elam, S., Rose, L., & Gallup, A. (1993). The 25th annual gallup/Phi Delta Kappa poll of the public's attitude toward public schools. *Phi Delta Kappan*, 75(2), 137-152.
- Futrell, M. (1989). Mission not accomplished: Education reform in retrospect. *Phi Delta Kappan*, 71(1), 8-14.
- Madaus, G. (1994). A technological and historical consideration of equity issues associated with proposals to change the nation's testing policy. *Harvard Educational Review*, 64(1), 76-95.
- Madaus, G., & Kellaghan, T. (1991). *Student examination systems in the European community: Lessons for the United States*. Washington, DC: Office of Technology Assessment.
- Metropolitan Life. (1993). *The Metropolitan Life survey of the American teacher 1992*. New York: Author.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for education reform*. Washington, DC: U.S. Government Printing Office.
- National Council on Education Standards and Testing. (1992). *Raising standards for American education*. Washington, DC: U.S. Government Printing Office.
- Office of Technology Assessment. (1992). *Testing in American schools. Asking the right questions*. Washington, DC: Congress of the United States.
- Popham, W. J. (1993). Circumventing the high costs of authentic assessment. *Phi Delta Kappan*, 74(6), 470-473.
- Pyke, N. (1993, May 7). Sharp rise in exclusions confirmed. *Time Educational Supplement*, p. 2.
- Shepard, L. (1991, April). *Effects of high-stakes testing in instruction*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Silcock, P. (1990). Implementing the national curriculum: Some teachers' dilemmas. *Education*, 3-13, 3-10.
- Smith, M., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 10(3), 7-11.