

The Fall and Rise of Standards-Based Education

By Robert J. Marzano and John S. Kendall

A National Association of State Boards of Education (NASBE) Issues in Brief

Preface

Over the last several years, state boards of education have become increasingly engaged in the development of content and performance standards and the assessments that will accompany them. In responding to board members' requests for additional information on this critical topic, The National Association of State Boards of Education (NASBE) asked Robert Marzano and John Kendall of the Mid-continent Research for Education and Learning (McREL), who have spent five years studying standards documents and related subject-area materials, to write an Issue Brief on the move toward standards-based education.

The resulting brief is divided into two main sections:

- 1) a **brief historical overview** that puts the standards movement in the context of recent education reform efforts and surveys some of the problems that have arisen in the development of standards at the national level; and
- 2) a **technical review** that discusses the nature and technical aspects of standards in more detail and summarizes the McREL resource data base that is available to states, districts and schools as an aide in their development of standards.

There are few issues more important to education policymakers today than the development of standards and assessments that fundamentally define what students should know and be able to do--and how these students should demonstrate their knowledge and skills. Yet first-rate standards at the state and local levels, as Marzano and Kendall and other reports have pointed out, will take years to develop and refine. I urge state board members to stay engaged in this process, know the issues, and ask the tough questions regarding their state's standards documents. These standards define--for students, teachers, parents, and the community--what your education system is all about.

Brenda Welburn
Executive Director
National Association of State Boards of Education

For more information, visit the [NASBE Web Site](#).

Part One

Many, if not most, educators are unaware of the impact the very discussion of standards has had on American education, let alone the reorganization of schools around standards. Educational policy analyst Anne Lewis (1995) writes that "whether lauded as a sign of progress or scorned as anathema," the standards movement is one of the most talked about issues in school reform (p. 745). Researchers Robert Glaser and Robert Linn assert that it might be only in retrospect that we recognize the importance of the current discussion of standards in American education:

In the recounting of our nations' drive toward educational reform, the last decade of this century will undoubtedly be identified as the time when a concentrated press for national education standards emerged. The press for standards was evidenced by the efforts of federal and state legislators, presidential and gubernatorial candidates, teacher and subject-matter specialists, councils, governmental agencies, and private foundations. (Glaser & Linn, 1993, p. xiii)

When and where did the discussion of standards originate? Where will it lead us? In this monograph, we attempt to answer these and other basic questions about standards.

A Brief History of the Modern Standards Movement

Former Assistant Secretary of Education, Diane Ravitch, is commonly recognized as one of the chief architects of the modern standards movement. In her book, *National Standards in American Education: A Citizen's Guide* (Ravitch, 1995), Ravitch explains the rationale for standards in a straightforward manner:

Americans...expect strict standards to govern construction of buildings, bridges, highways, and tunnels; shoddy work would put lives at risk. They expect stringent standards to protect their drinking water, the food they eat, and the air they breathe....Standards are created because they improve the activity of life. (pp. 89)

Ravitch asserts that just as standards improve the daily lives of Americans, so too will they improve the effectiveness of American education: "Standards can improve achievement by clearly defining what is to be taught and what kind of performance is expected." (p. 25)

Many educators see the publication of the now famous report, *A Nation at Risk*, as the initiating event of the modern standards movement. Ramsay Seldon, Director of the State Assessment Center at the Council of Chief State School Officers, notes that after this prominent expose on public education, state and local leaders set out to improve the education system through new policies such as increasing the rigor of graduation requirements. When these efforts produced disappointing results, policymakers turned to national goals and standards:

[T]here was a feeling of urgency that the education system needed to be stronger, and that in addition to what states and districts and individual schools were doing we needed a stronger presence at the national level....We recognized that we didn't need a national curriculum, so national goals and voluntary national

standards came to be seen as a good mechanism for providing a focus. (In O'Neil, 1995, p. 12)

Researcher Lorrie Shepard also cites *A Nation at Risk* as a critical factor in the modern standards movement. Shepard (1993) notes that after the publication of the report, the rhetoric of educational reform changed drastically. It began to make a close link between the financial security and economic competitiveness of the nation and our educational system. Who will soon forget the chilling words often quoted from *A Nation at Risk*: "The educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a nation and a people....We have, in effect been committing an act of unthinking, unilateral educational disarmament." (National Commission on Excellence in Education, 1983, p. 5)

These growing concerns about the educational preparation of the nation's youth prompted President Bush and the nation's governors to call an Education Summit in Charlottesville, Virginia in September, 1989. At this summit, President Bush and the nation's governors, including then-governor Bill Clinton, agreed on six broad goals for education to be reached by the year 2000. Two of those goals (3 and 4) related specifically to academic achievement:

Goal 3: By the year 2000, American students will leave grades 4, 8, and 12 having demonstrated competency in challenging subject matter including English, mathematics, science, history, and geography; and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our modern economy.

Goal 4: By the year 2000, U.S. students will be first in the world in science and mathematics achievement. (National Education Goals Panel, 1991, p. ix)

Soon after the summit, two groups were established to implement the new educational goals: the National Education Goals Panel (NEGP) and the National Council on Education Standards and Testing (NCEST). Together, these two groups were charged with addressing unprecedented questions regarding American education such as: What is the subject matter to be addressed? What types of assessments should be used? What standards of performance should be set?

The summit and its aftermath engendered a flurry of activity from national subject matter organizations to establish standards in their respective areas. Many of these groups looked for guidance from the National Council of Teachers of Mathematics who preempted the public mandate for standards by publishing the *Curriculum and Evaluation Standards for School Mathematics* in 1989. As education reporter Karen Diegmueller (1995) explains, the NCTM standards "redefined the study of math so that topics and concepts would be introduced at an earlier age, and students would view math as a relevant problem-solving discipline rather than as a set of obscure formulas to be memorized." (Diegmueller, 1995, p. 5) The National Academy of Sciences used the apparent success of the NCTM standards as the impetus for urging Secretary of Education Lamar Alexander to underwrite national standards-setting efforts in other content areas. According to Diane Ravitch, then an assistant secretary of education, "Alexander bankrolled the projects out of his office's discretionary budget." (Diegmueller, 1995, p. 5) The National Science Teachers Association and the American Association

for the Advancement of Science quickly launched independent attempts to identify standards in science. Efforts soon followed in the fields of civics, dance, theater, music, art, language arts, history, and social studies, to name a few.

Troubled Times Despite the publicity given the national goals, the federal support for standards, and the enthusiasm of educators from various subject areas, critics of the standards movement also caught the public's attention. Among the issues raised by these critics were:

Resources--Some saw the standards movement as a major drain on resources that should be used for more pressing needs such as basic educational materials. For example, TheodoreSizer, founder of the Coalition of Essential Schools, noted that in many classrooms, "The maps on the walls still call [Zaire] the Belgian Congo. Those are the things that just cry out for attention." (In Diegmueller, 1995, p. 5)

Educational Apartheid--Other critics saw the standards movement as another burden that would be placed on the shoulders of those who traditionally do not do well in schools. Curriculum professor Michael Apple noted that: National standards and national testing are the first steps toward educational apartheid under the rhetoric of accountability. (Diegmueller, 1995, pp. 56)

Standards as New Attempts at Previous Failed Reforms--Still others saw the standards movement as a thinly veiled attempt at a type of educational reform that had been tried a number of times before. For example, Elliot Eisner noted the similarity of the standards movement to the efficiency movement of the early 1900s:

The efficiency movement, which began in 1913 and lasted until the early 1930s, was designed to apply the principles of scientific management to schools. Its progenitor, Frederick Taylor, the inventor of time-and-motion study, was a management consultant hired by industrialists to make their plants more efficient and, hence, more profitable. By specifying in detail the desired outcomes of a worker's efforts and by eliminating "wasted motion", output would increase, profits would soar, wages would rise, and everyone would benefit. (Eisner, 1995, p. 159)

According to Eisner, school administrators soon found that the basic concept underlying the efficiency movement--namely that one could mechanize and routinize teaching and learning--did not work. Educators would no doubt come to the same conclusions about standards, opined Eisner.

The standards movement was also likened to the failed behavioral objectives movement of the 1960s. Again, the basic notion behind behavioral objectives was to define educational goals in terms that were sufficiently specific to determine without ambiguity whether or not students had achieved them. Through his book, *Preparing Instructional Objectives*, published in 1962, Robert Mager routinized and popularized the process of constructing behavioral objectives to such an extent that teachers in virtually every subject area, at every grade level, were writing behavioral objectives all across the country during the 1960s. For Mager, an objective must identify the expected behavior in detail, the conditions in which it is to be displayed, and the criterion that make it possible to measure the student's performance in relation to the criterion. An example of a behavioral objective following Mager's criteria would be: "At the end of a 50-minute

period of instruction, students will be able to complete eight out of ten problems in two-column addition within a five minute period."

This level of detail, although possibly effective instructionally, created a system that was overwhelming for teachers. As Eisner notes, the approach required that schools construct hundreds and sometimes thousands of behavioral objectives to specify the outcomes of instruction. Soon, schools and districts became bogged down by the sheer weight of numbers, and the movement lost steam.

Content--In addition to its association with the flawed efficiency and behavioral objective movements of the past, the standards movement received a fair amount of criticism for the very content it promoted. Perhaps the lowest point in the standards movement was the debate over the history standards. In the fall of 1994, Lynne V. Cheney, a fellow of the American Enterprise Institute, unleashed an attack on the U.S. History standards that, along with science, had been the first standards project to receive funding from the Department of Education in 1991. Cheney accused the history standards of portraying the United States and its white, male-dominated powerstructure as an oppressive society that victimizes minorities and women. She further charged that the history standards ignored such traditional historical figures as George Washington and Robert E. Lee to placate proponents of multi-culturalism. Suddenly, the rather academic discussion of standards burst onto the national scene. Diegmueller notes that:

Cheney's views won such exceptionally wide exposure because, as chairwoman of the National Endowment for the Humanities, she had lobbied for history standards, funded the project, and selected its leaders and many of the people on its 29-member board. Soon it became evident that the criticism was not about to subside--even though there were far more supporters than detractors. The U.S. Senate even weighed in, denouncing the history standards by a vote of 99 to 1. (Diegmueller, p. 8)

To date, the history standards have not recovered from the negative public perception generated by Cheney's criticisms.

Volume of Material--Perhaps the ultimate criticism of the national efforts to establish standards was the charge that, once developed, they were simply too cumbersome to use. In the beginning, policymakers and educators had expected to see concise standards that were symmetrical in tone and format. However, as the standards drafts and final documents were produced, it became clear that they were far from concise. Chester Finn noted that: "The professional associations, without exception, lacked discipline. They all demonstrated gluttonous and imperialistic tendencies." (Diegmueller, p. 6)

At the time of Finn's statement in 1995, the standards documents, taken together, weighed about 14 pounds, stood six inches tall and contained over 2,000 pages. Since then, more documents, more pounds, and more inches have been added to the total mass of standards. By contrast, the Japanese national curriculum fits into "three slender volumes, one for elementary, one for lower secondary, and one for upper secondary." (Ravitch, 1995, p. 15) Ron Brandt, Executive Editor of the Association for Supervision and Curriculum Development, acknowledged the problem of the sheer volume of the standards in the following way:

I would describe them as an ambitious conception of what professional educators, most of whom are advocates or specialists in the various school subjects, want students to learn in those subjects. It's the classic curriculum dilemma faced by every principal, central administrator, and generalist teacher: specialists naturally expect a lot; they love their subject and they know its possibilities. Taken as a whole, however, such statements of aspirations are overwhelming. (Brandt, 1995, p. 5)

In summary, the once bright promise of subject area standards, born from a desire to improve the rigor and effectiveness of American education, has faded under a wide array of criticisms, and the movement itself has become bogged down under its own weight.

Is the Standards Movement Still Alive?

Given the intense criticism of many aspects of the modern standards movement, there are some who believe that it is, for all practical purposes, dead. Ron Brandt explains:

Now that some of the original sponsors are disappointed in the new standards because they are not what was expected, what does that mean for educators? Apparently, these standards will not soon become a national curriculum or the basis for a set of high-stakes tests. Under the circumstances, educators can breathe a sigh of relief and, with discretion, put them to use in the endless task of improving curriculum and instruction. (1995, p. 5)

In general, we agree with Brandt that America will not soon have a set of nationally accepted standards. In addition to the problems cited above with the standards as developed by the national subject matter groups, the impetus for reform at the federal level has been halted because of a changing political climate. This has been dramatically illustrated by the demise of the National Education Standards and Improvement Council (NESIC). Created as part of the Goals 2000 legislation passed in 1994, NESIC was supposed to oversee the development of voluntary national content standards and "certify" the standards created by states. But by June of 1995, education policy analyst David Cohen was writing that "NESIC seems to be dead on arrival. Barely half a year after Goals 2000 was signed into law, Republicans took control of Congress. Although many Republicans had supported the legislation in the previous Congress, the new faces were generally more conservative and had little use for any sort of national school reform. They had especially little use for an agency that would devise, promulgate and certify national educational standards." (Cohen, 1995, p. 752)

At the same time, the standards movement at the state level has also been problematic. Campaigns have been mounted to stop the identification of state standards in Virginia, Colorado, Oregon, Pennsylvania and Washington, to name a few. A recent study by the American Federation of Teachers (Gandal, 1995) found that state standards are, for the most part, weak: Only 13 states have standards that are strong enough to carry the weight of the reforms being built upon them. (p. 13)

However, **we do not believe that the standards movement is dead.** In fact, we assert that the logic behind organizing schooling around standards is so compelling as to make standards-based school reform something that schools and districts will implement even in the absence of federal or state mandates or incentives. Indications are that the

standards movement, though fallen from grace at the national level, is rising in reform efforts at the local level; over the last year, the professional development arm of our organization, the McREL Institute, has seen a greater than threefold increase in the number of districts and schools who have contracted for assistance in the development of standards and benchmarks. And even the AFT study concluded that it is not too late "in most states for changes to be made that will strengthen their standards and enhance their efficacy in improving student achievement." (p. 31) There appear to be at least four reasons why standards represent one of the most powerful options for school reform:

Reason #1: The Erosion of the Carnegie Unit and the Common Curriculum

Although 90 years old, the Carnegie unit is still a basic structural feature of American education. As initially defined in 1906 by the President of the Carnegie Foundation for the Advancement of Teaching, the Carnegie unit was "a course of five periods weekly throughout an academic year." (In Tyack and Tobin, 1994) By convention, these periods had come to be thought of as 55 minutes long. A committee convened by the Carnegie Foundation also set standards for the content and duration of specific courses, "specifying in great detail the content of units in subjects like English, mathematics, Latin, Greek, foreign languages, history and science. Thus, they standardized not only time and credits, but gave pride of place to traditional academic subjects...." (p. 461).

Initially, then, the Carnegie unit represented an implicit set of standards. As adopted by public school systems, Carnegie units required that high schools cover specified content in a specified period of time. For decades, this system worked fairly well.

Yet over the years schools have moved away from a central core of knowledge and skill. From the 1940s until the mid-1970s, the emphasis on serving the interests of individual children generated a geometric expansion of the number of courses that constituted the high school curriculum. By the mid-1970s, the U.S. Office of Education reported that more than 2,100 different courses were being offered in American high schools. (Ravitch, 1995, p. 37)

This trend toward ever-expanding offerings and ever-decreasing uniformity in the school experience still exists today. This is evident in studies that have focused on how teachers use time. To illustrate, in a study of the content teachers emphasize within reading and the language arts, Berliner (1984) found that one fifth grade teacher could find only 68 minutes a day of instruction in reading and language arts, while another teacher was able to find 137 minutes a day. At the second grade level one teacher allocated 47 minutes a day for reading and language arts, while another teacher managed to find 118 minutes a day, or 2 1/2 times more per day to teach reading and language arts.

In summary, where the content covered and the manner in which time is spent was at one time fairly uniform in American education, today there is little consistency in how much time students spend on a given subject or the knowledge and skills covered within that subject area.

Reason # 2: The Variation in Current Grading Practices

Most educators and non-educators assume that grades are precise indicators of what students know and can do with a subject area. In addition, most people assume that

current grading practices are the result of a careful study of the most effective ways of reporting achievement and progress. In fact, current grading practices developed in a fairly serendipitous way.

Mark Durm (1993) provides a detailed description of the history of grading practices in America, beginning in the 1780s when Yale University first started using a four-point scale. By 1897, Mount Holyoke College began using the letter grade system that is so widely used in education today.

For the most part, this 100-year-old system is still in place today. Unfortunately, even though the system has been in place for a century, there is still not much agreement as to the exact meaning of letter grades. This was rather dramatically illustrated in a nationwide study by Robinson & Craver (1988) that involved over 800 school districts randomly drawn from the 11,305 school districts with 300 or more students. One of their major conclusions was that districts stress different elements in their grades. While all districts include academic achievement, they also include other significant elements such as effort, behavior, and attendance. Figure 1 lists the percentages of districts that include each of these variables in their grades.

Figure 1. Percentage of Districts Reporting Use of Effort, Behavior, and Attendance in Determining Grades

Grade Level	Effort	Behavior	Attendance
K	26%	4%	6%
1-3	26%	4%	7%
4-6	26%	4%	7%
7-9	32%	7%	14%
10-12	33%	8%	17%

The Robinson & Carver study was done using the official policies of school districts. In a separate study, we polled individual teachers on the extent to which they include effort, behavior, attendance, and the added variable, cooperation, in their grades. These findings are shown in Figure 2:

Figure 2. Percentage of Teachers Reporting Use of Effort, Behavior, Cooperation, and Attendance in Determining Grades

Grade Level	Effort	Behavior	Cooperation	Attendance
K (n=79)	31%	7%	4%	8%
1-3 (n=110)	29%	8%	4%	8%
4-6 (n=158)	30%	8%	8%	10%
7-9 (n=142)	36%	10%	8%	18%
10-12 (n=151)	36%	14%	9%	24%

Figures 1 and 2 imply that there is great discrepancy in the factors teachers consider when they construct grades. In effect, we have a situation in which grades given by one teacher might mean something entirely different from grades given by another teacher

even though the teachers are presiding over two identical classes with identical students who do identical work. Where one teacher might count effort and cooperation as 25% of a grade, another teacher might not count these variables at all.

Reason #3: The Lack of Attention to Educational Outputs

Perhaps the most compelling argument for organizing educational reform around standards is the shift in emphasis from what schools put into the process of schooling to what we get out of schools that is, a shift from educational "inputs" to educational "outputs". Chester Finn describes this shift in perspective in terms of an emerging paradigm for education:

Under the *old* conception...education was thought of as process and system, effort and intention, investment and hope. To improve education meant to try harder, to engage in more activity, to magnify one's plans, to give people more services, and to become more efficient in delivering them.

Under the *new* definition, now struggling to be born, education is the result achieved, the learning that takes root when the process has been effective. *Only* if the process succeeds and learning occurs will we say that *education* happened. Absent evidence of such a result, there is no education--however many attempts have been made, resources deployed, or energies expended. (Finn, 1990, p. 586)

Finn asserts that the shortcoming of the old input paradigm of schooling came to light in the mid-1960s when the country set out to provide disadvantaged and minority students with better opportunities by providing them with better education. The U.S. Office of Education was commissioned by Congress to conduct a major study of the quality of educational opportunity. The result was the celebrated "Coleman Report" (after chief author and researcher, James Coleman), which was released in 1966. Finn explains that the report concluded that input variables might not actually have all that much to do with educational equality when equality was conceived of in terms of what students actually learned as opposed to the time, money, and energy that were expended. In later years, Coleman wrote about the study that its

major impact [was] in shifting policy attention from its traditional focus on comparison of inputs (the traditional measures of school quality used by school administrators: per-pupil expenditures, class size, teacher salaries, age of building and equipment, and so on) to a focus on output, and the effectiveness of inputs for bringing about changes in output. (Coleman, 1972, pp. 149-150)

According to Finn, while many school reform efforts are still grounded in the old paradigm, some are beginning to embrace the output view of accountability. Among these new efforts, Finn cites the national goals established at the education summit in 1989:

Perhaps even more portentous was the 1989 "education summit" held in Charlottesville, Virginia, at which the nation's governors and President Bush actually agreed to develop a set of national "goals" for education--goals that, as they were hammered out and made public in early 1990, have far more to do with

outcomes than with service delivery. They also pledged to issue annual "report cards" on progress toward those goals. (Finn, 1990, p. 591)

In summary, the new, more efficient and accountable view of education is output-based-- outputs defined in terms of specific student learnings in terms of specific standards.

Reason #4: Other Countries Do It

A final reason for considering a standards-based approach is that most of the other countries we say that we want to emulate rely on policies and structures that are fundamentally standards-based in nature. For example, in their study of standards-setting efforts in other countries, Resnick and Nolan (1995) note that Many countries whose schools have achieved academic excellence have a national curriculum. "Many educators maintain that a single curriculum naturally leads to high performance, but the fact that the United States values local control of schools precludes such a national curriculum." (p. 9)

Although they caution that a well articulated national curriculum is not a guarantee of high academic achievement, Resnick and Nolan offer some powerful illustrations of the effectiveness of identifying academic standards and aligning curriculum and assessments with those standards. France is a particularly salient example:

In texts and exams, the influence of the national curriculum is obvious. For example, a French math text for 16-year-olds begins by spelling out the national curriculum for the year so that all 16-year-olds know what they are expected to study. The book's similar table of contents shows that the text developers referred to the curriculum. Moreover, the text makes frequent references to math exams the regional school districts have given in the past. Students practice on these exams to help them prepare for the exam they will face; they know where to concentrate to meet the standard. (p. 9)

In a similar vein, a report published by NESIC, the National Education Standards and Improvement Council (1993), details the highly centralized manner in which standards are established in other countries. For example, in China, standards are set for the entire country and for all levels of the school system by the State Education Commission in Beijing. In England, standard setting was considered the responsibility of local schools until 1988, when the Education Reform Act mandated and outlined the process for establishing a national curriculum. The School Examinations and Assessment Council was established to carry out this process. In Japan, the ministry of education in Tokyo (Manibushi) sets the standards for schools, but allows each of the 47 prefectures (Ken) some latitude in adapting those standards. According to the NESIC report, "Most countries embody their content standards in curriculum guides issued by the ministries of education or their equivalents." (pc-51) Additionally, "A national examination system provides a further mechanism for setting standards through specifications of examinations, syllabuses and regulations, preparations of tests, grading of answers, and establishment of cutoff points." (pc-51)

What Have We Learned?

From history and from our work with schools and districts, we have learned at least two very important lessons about organizing schools around standards. First, **setting standards is a very technical process that should not be taken lightly**. A state, school or district that sets content area standards by convening group of teachers and administrators to identify standards without serious up-front consideration of technical issues, is asking for trouble. Second, **organizing schooling around standards is not a cookie cutter process--no one size fits all**. Standards-based approaches must be tailor made to the specific needs and values of individual schools and districts. In Part Two we consider some technical issues that surround the setting of standards.

Part Two

Technical Issues and the McREL Data Base

A school or district wishing to construct local standards, or augment their state standards, has a great deal of work ahead of it. Relative to the latter situation--augmenting state standards--the American Federation of Teachers report cited earlier (Gandal, 1995), found that at this point the vast majority of state documents are not specific enough for schools and districts to use to restructure their curriculums. Additionally, the vast majority of state documents deal with only three levels (e.g., grades K-5, 6-8, 9-12) or four levels (e.g., grades K-2, 3-5, 6-8, 9-12), yet most districts must deal with standards at all grade levels. Finally, many state documents are presented as "guidelines" to be used by local districts rather than as mandated standards that must be followed without alteration. Consequently, even where a sound state standards document exists, the schools and districts within that state will still have to do a great deal of standards writing and redesign if they wish to implement a standards-based approach.

As an aid in their design work, schools or districts (or states) might turn to the "national standards documents" from the various subject specific organizations. For example, the National Council of Teachers of Mathematics (NCTM) has published the *Curriculum and Evaluation Standards for School Mathematics* (1989), and the American Association for the Advancement of Science (AAAS) has published the *Benchmarks for Science Literacy* (1993). While one might assume that schools and districts need only consult these subject specific documents and copy standards verbatim to construct their local standards, this is not the case. This is because the various national standards documents vary conceptually in a number of important ways. For example, the manner in which standards are described or defined by a document in mathematics, let's say, might be quite different from the manner in which standards are defined within a document that focuses on science. Additionally, some content areas have multiple documents that identify standards and each document might take a slightly different or greatly different perspective on standards. For example, mathematics standards are identified within the NCTM standards document. However, mathematics standards are

also identified in the following documents, each of which has slight and sometimes great differences in the ways in which standards are described:

Benchmarks for Science Literacy by the American Association for the Advancement of Science, 1993.

Mathematics Assessment Framework by the National Assessment of Educational Progress, 1992.

What Work Requires of Schools: A SCANS Report for America 2000 by the Secretary's Commission on achieving Necessary Skills, 1991.

Workplace Basics: The Essential Skills Employers Want by Carnevale, Gaines and Meltzer, 1990.

In short, a state, school or district wishing to establish standards based on the national documents must first identify what they mean by a standard and the format their standards will take. Next, they must systematically analyze all the national documents translating them into a format and conceptual base compatible with their own. This, of course, can be a labor-intensive endeavor that requires valuable resources of a variety of types.

Our work with the McREL data base provides a convenient entree for explaining some of the technical issues about standards that policymakers and others working in education reform need to understand. As we began our efforts, we found that there were at least four key issues on which we had to take a very clear position--the issues of: 1) content versus curriculum standards; 2) types of content standards; 3) content versus performance standards; and 4) the need for levels of standards. Indeed, being aware of these four issues is key for anyone engaged in the standards movement. We briefly consider each of them here. For a detailed discussion of these and other issues, see *Content Knowledge: A Compendium of Standards and Benchmarks for K12 Education*.

1. Content Versus Curriculum Standards

A number of documents we analyzed contained a mixture of content standards and curriculum standards, yet did not make a distinction between the two. In simple terms, a **content** standard describes what students should know and be able to do; a **curriculum** standard describes what should take place in the classroom. Specifically, curriculum standards address instructional technique or recommended activities as opposed to knowledge and skill per se. The difference between the two can be illustrated from the following two statements from the National Council of Teachers of Mathematics (NCTM, 1989) framework. Within the document both statements are presented as elements of standards:

- a. *use estimation to check the reasonableness of results*
- b. *describe, model, draw and classify shapes*

Element a) describes a skill or ability a person might use to solve a real life problem. For example, you might use estimation to check the reasonableness of your calculations as to the amount of wood you would need to build a fence around your back yard. Element b), on the other hand, is not commonly used in real life situations. That is, it is difficult to

imagine many situations that would demand the skill of being able to model, draw, or classify shapes, whether to solve an academic or a day-to-day problem. Rather, this kind of activity is best described as an instructional device to help students understand shapes or to provide a way for them to demonstrate their understanding of shapes. Therefore, it is a **curriculum** standard rather than a **content** standard. It might be said that curriculum standards describe the methods used to help students achieve content standards.

The McREL data base has content standards as its focus. There are two overarching reasons for this choice. First, content standards describe the goals for individual student achievement while curriculum standards provide information that contributes to reaching these goals. Second, curriculum standards, which usually focus on activities, projects or techniques, if interpreted rigidly, could leave teachers with little or no room for instructional diversity. It is important to note that even though we have not included curriculum standards in our data base, we assume that a school or district might wish to generate curriculum standards. Given that content standards have been clearly articulated, the complementary set of curriculum standards could be viewed as useful instructional suggestions for accomplishing the learning goals implicit in the content standards.

2. Types of Content Standards

Since our approach has a content (as opposed to curricular) orientation, the standards we have identified assume some of the characteristics of content area knowledge. Specifically, the standards identified within our data base generally fall into three broad categories representing the three general types of knowledge as exemplified in Figure 3:

Figure 3: Types of Standards

Procedural	Declarative	Contextual
performing long division	understanding the concept of a numerator	modeling numbers using number line
setting up an experiment	knowing what an amoeba is	classifying organisms
editing an essay	knowing the conventions of punctuation	using appropriate tone and style for a selected audience

In Figure 3, the first column provides examples of procedural knowledge, the skills and processes important to a given content area. The examples in the second column involve declarative knowledge. Declarative knowledge can be thought of as "information" and usually involves component parts. For example, knowledge of the concept of "democracy" includes understanding that decisions are made by the people, each person has a single vote, votes are weighted equally, and so on.

The last column contains items that are not simply declarative or procedural, but specify knowledge in context, information and/or skills that have particular meaning because of the conditions that form part of their description. Like the declarative/procedural distinction, this contextual knowledge is basic; a "piece" that cannot be further reduced

without loss of important information. For example, modeling numbers using a number line involves a procedural part (the process of modeling) and a declarative part (the concept of numbers). However, the two combined are greater than the sum of the individual parts. The combination represents a basic unit of knowledge important to the domain of mathematics. The process of modeling in this context has specific characteristics that it does not have in other contexts and the characteristics of numbers that are highlighted in the modeling process are probably not highlighted quite so specifically in any other environment.

Given the unique features of procedural, declarative and contextual knowledge, we have chosen to code the various component parts of our data base as to which type of knowledge they represent. This is not to say that all standards developers would attend to this level of detail. However, given that the purpose of the resource data base is to create a flexible tool to be used by states, schools, and districts, we believed that such specificity was important, if for no other reason than to provide educators with a sense of the differing "structures" of various content areas. For example, based on the documents we have analyzed thus far, it appears that mathematics knowledge is 46% declarative, 41% procedural, and 13% contextual. However, science knowledge is 95% declarative, 2.5% procedural, and 2.5% contextual. Those constructing or reviewing their own standards using our data base might want to preserve these proportions or consciously emphasize one type of knowledge over another due to local priorities. For example, a school or district might want to emphasize procedural and contextual scientific knowledge, which deals with the scientific process, in lieu of some of the declarative scientific knowledge which deals more with scientific facts and concepts.

3. Content or Performance Standards?

One of the significant controversies within the developing science of standards-based education is whether standards should be content or performance based. Some take a clear content position describing standards in terms of knowledge and skill that should be acquired; others describe standards in terms of tasks through which students demonstrate knowledge and skill by their performance. This is the performance position. Performance standards differ from curriculum standards in that, like content standards, they are not designed as activities for the sake of instruction, but rather are descriptions, via tasks, of what it is students should know and be able to do to demonstrate competence.

The content position focuses on clearly defined declarative, procedural or contextual knowledge. The performance position presumes this knowledge is defined if it is embedded in a task, even though this task must be a narrower application of the knowledge. In other words, a content standard is a **statement of the knowledge** or understanding we would expect students to have. On the other hand, a performance task **describes a specific use of knowledge and skills**: it is not a description of knowledge, but a description of some application of it. For example, a content standard in science might specify that students should understand the characteristics of ecosystems on the earth's surface. The performance standard for that piece of declarative knowledge would specify the level of accuracy and the facts, concepts and generalization about ecosystems on the earth's surface that a student must understand to be judged as having obtained a suitable level of achievement. It would also put that knowledge in some type of performance environment by stating that the information must be presented, for example, in the form of an essay, a simulation, or an oral report

with accompanying graphics. As the National Education Standards and Improvement Council notes:

...performance standards indicate "both the nature of the evidence (such as an essay, mathematical proof, scientific experiment, project exam, or combination of these) required to demonstrate that content standards have been met and the quality of student performance that will be deemed acceptable..." (NESIC, 1993, p. 22)

We believe that performance standards are a critical component of a comprehensive, standards-based approach to schooling. In fact, performance standards and content standards have a hand-in-glove relationship. As Marc Tucker, co-director of the New Standards Project, has noted: "You can't assess kids' performance unless you give them a task, and you can't assess their degree of achievement unless they actually perform the task." (Tucker, 1992, p. S.3)

Content standards, then, can be effectively translated into assessable elements via the articulation of performance standards. We, therefore, **recommend that schools and districts articulate a set of content standards and a complementary set of performance standards**. The content standards identify what students should know and be able to do. The performance standards identify the environments in which that knowledge and skill should be demonstrated. Unfortunately, we see a trend nationally for schools and districts to develop performance standards only. Presumably, these schools and districts assume that a well-articulated set of performance standards will implicitly include content standards. There are two problems with such an approach.

First, by virtue of limiting the expression of knowledge and skill to that demonstrated in a particular task, performance standards have a limited scope, and it could take many performance standards to exhaust, if possible, the potential applications for any set of given content knowledge. Second, and still more problematic from the point of view of covering important knowledge and skills, performance tasks in isolation are rarely transparent as to the knowledge and skills required for their successful completion. To illustrate, consider the example of an open-ended performance task used in one state's assessment: "How much does it cost to take a shower?" (Wiggins, 1993, p. 204) To ask students to determine the cost of a shower is an excellent, real-world challenge; but if, in this state, the performance tasks are not based on specific content standards, it is not immediately evident what declarative, procedural or contextual knowledge this task is designed to assess. Initially, it might seem that an understanding of the British Thermal Unit (BTU) is the declarative knowledge critical to the "shower" task. However, without the explicit guidance given by a set of content standards, we must make a calculated guess that this is the critical knowledge intended as the focus of the task. On the other hand, if we have already determined through content standards that students should understand BTUs, there are any number of tasks (including the shower task) that we could construct to confirm whether the student has this knowledge, and how well the student knows it.

We believe that performance tasks (as standards) make clear how knowledge and skills are useful, but tasks in themselves fail to make clear all that is important. Therefore, states, schools, and districts would be advised to begin with content standards, and then use them to generate a complementary set of performance tasks.

4. The Need for Levels of Standards

Even a cursory review of the standards generated by different groups reveals very different perspectives on the level of generality at which standards should be stated. For example, the National Standards for Arts Education (1994, p. 34) provides this as a standard:

Understand[s] the arts in relation to history and cultures

In contrast, a draft document from the National History Standards Project (1994, p. 84) lists the following as a standard:

Know[s] the causes of the Civil War

The example from the National History Standards Project is obviously more specific than that from the National Standards for Arts Education. In addition, the History document provides a much more detailed level of sub-component information for its standards than does the Arts document. The extent to which standards are articulated in *general* versus *specific* terms is critical since the level of generality adopted by a state, school or district will affect the level of detail within the standards, the kind of comprehensiveness the standards aim for, and the number of standards produced.

The approach we have adopted is to articulate standards at a general level, yet define specific subcomponents at various developmental levels. These developmentally appropriate subcomponents are referred to as "**benchmarks.**" To illustrate, consider the following content standard within mathematics: "*demonstrates number sense and an understanding of number theory.*" This statement maps out a very general area within mathematics. Benchmarks for this standard appropriate at the high school level might include the following:

- Understands characteristics of the real number system and its subsystems.
- Understands the relationship between roots and exponents.
- Models numbers using three-dimensional regions.

Benchmarks appropriate for middle school would include the following:

- Understands the relationship of decimals to whole numbers.
- Understands the relationship of fractions to decimals and whole numbers.
- Understands the basic difference between odd versus even numbers.
- Understands the basic characteristics of mixed numbers.
- Models numbers using number lines.

Benchmarks, then, describe the specific developmental components of the general domain identified by a standard. Within our data base, benchmarks are provided at four levels, roughly corresponding to grades K-2 (Level I), 3-5 (Level II), 6-8 (Level III), and 9-12 (Level IV). The benchmarks within a given standard are intended as expectations for the upper end of the interval in which they are presented. To illustrate, below are the science benchmarks listed for the interval K-2 within the standard, "*Understands essential ideas about the composition and structure of the universe and the Earth's place in it.*"

- Knows that the stars are innumerable, unevenly dispersed, and of unequal brightness
- Knows that the sun can be seen only in the day time, whereas the moon is out sometimes at night and sometimes during the day
- Knows that the moon looks a little different every day, but looks the same again every four weeks

(Kendall & Marzano, 1996, pp. 79-80)

These are expectations of what second graders should know. To translate these benchmarks into grade-level benchmarks, a school or district need only "map backwards" identifying which elements would be deleted or altered at lower grade levels. For example, first grade teachers might decide that the benchmarks above should be restated in the following way to be developmentally appropriate for first grade students:

- Knows that there are many stars scattered all over the sky
- Knows that the sun is seen during the day and the moon can be seen at night and sometimes
- during the day
- Knows that the moon changes in appearance

Similarly, kindergarten teachers might conclude that the benchmarks should be written in the following way to be developmentally appropriate for kindergarten students:

- Knows that stars are seen in the sky at night
- Knows the sun is seen during the day and the moon is usually seen at night

Thus, benchmarks written generally at four levels (as in our data base) can easily be expanded to individual grade levels.

The Format of the McREL Data Base

In all, the McREL data base contains 201 different standards and their related benchmarks. These standards are organized into thirteen major categories as follows:

- **Mathematics:** 9 standards, 349 benchmarks
- **Science:** 18 standards, 324 benchmarks
- **History:**
 - K-4 History: 4 standards, 55 benchmarks;
 - U.S. History: 10 standards, 141 benchmarks;
 - World History: 13 standards, 157 benchmarks;
- **Historical Understanding:** 2 standards, 42 benchmarks
- **Language Arts:** 13 standards, 372 benchmarks
- **Geography:** 18 standards, 238 benchmarks

- **Arts:**
 -
 - Dance: 6 standards, 62 benchmarks
 -
 - Music: 7 standards, 80 benchmarks
 -
 - Theatre: 6 standards, 72 benchmarks
 -
 - Visual Arts: 5 standards, 42 benchmarks
- **Art Connections:** 1 standard, 13 benchmarks
- **Civics:** 29 standards, 427 benchmarks
- **Economics:** 10 standards, 173 benchmarks
- **Foreign Language:** 5 standards, 86 benchmarks
- **Health:** 10 standards, 136 benchmarks
- **Physical Education:** 5 standards, 105 benchmarks
- **Behavioral Studies:** 4 standards, 100 benchmarks
- **Life Skills:**
 -
 - Thinking & Reasoning: 6 standards, 117 benchmarks
 -
 - Working with Others: 5 standards, 51 benchmarks
 -
 - Self-regulation: 6 standards, 59 benchmarks
 -
 - Life Work: 9 standards, 90 benchmarks

These standards were constructed from the content of 85 documents reported in the Appendix. These documents range from nationally funded efforts such as the history standards developed by the National Center for History in the Schools, to state documents such as the California Department of Education science framework, to documents developed through privately funded efforts used as elementary and junior high school standards developed by the Edison Project. Quite obviously, it would take a school or district an inordinate amount of time to analyze the 85 documents listed in the Appendix at the level of detail we have undertaken. However, using the McREL data base, standards developers can identify declarative, procedural, and contextual benchmarks and the national documents in which those benchmarks are implicitly or explicitly stated, as well as identify the interrelationship between benchmark elements. A school or district should then be able to construct their own standards and benchmarks and an accompanying set of performance tasks.

Tailoring a Standards-Based Approach to Local Needs

Ultimately, states, schools, and districts must design standards to meet the needs of their community. In our experience, we have found that this tailoring process is tantamount to answering a series of specific questions. The full complement of questions standards developers must address and their possible answers are described in depth in the book *How to Design a Standards-Based District, School or Classroom* (Marzano & Kendall, in press). Here we consider four of the many questions articulated in that book:

1. How many standards and benchmarks will be articulated?

In all, the McREL data base lists 201 standards and 3,291 benchmarks for implementation in K-12 schooling. Clearly, a school or district could not expect a student to demonstrate competence in all of these (although they may be a part of instruction); sheer numbers would make such a system untenable. Given that there are 180 days in the school year and 13 years of schooling (assuming students go to kindergarten), there are only 2,340 school days available to students. If all benchmarks in the McREL data base were addressed, this would mean that students would have to learn and demonstrate mastery in one or more benchmarks every school day, or more than seven benchmarks every week.

Thus, a school or district will surely have to select from the standards and benchmarks presented in the database if it wishes to construct a system in which students are to be held accountable for each benchmark. A reasonable number of benchmarks seems to be about 600, distributed in roughly the following way:

Level I: K-2: 75
Level II: 3-5: 125
Level III: 6-8: 150
Level IV: 9-12: 250

Quite obviously, to implement this 600-benchmark cap, schools and districts would have to exclude quite a few of the benchmarks currently identified.

2. Will all selected benchmarks be considered necessary to demonstrate competence in a standard?

One possible way to alleviate the problem of too many benchmarks is to consider benchmarks as exemplars rather than as necessary components of a standard. Using this option, students would be held accountable for demonstrating a mastery of a sample of the benchmarks within a level for a given standard as opposed to all the benchmarks within a given level.

To illustrate, consider the benchmarks in Figure 4 for the science standard "Understands energy types, sources, and conversions, and their relationship to heat and temperature."

Figure 4. Benchmarks for Model Science Standard

Level I (Grades K-2)

- Knows that the sun applies heat and light to earth
- Knows that heat can be produced in many ways (e.g., burning, rubbing, mixing chemicals)
- Knows that electricity in circuits can produce light, heat, sound, and magnetic effects

Level II (Grades 3-5)

- Knows that things that give off light often also give off heat
- Knows that mechanical and electrical machines give off heat
- Knows that heat can move from one object to another by conduction
- Knows that some materials conduct heat better than others; materials that do not conduct
- heat well can reduce heat loss
- Knows that electrical circuits require a complete loop through which the electrical current can
- pass

Level III (Grades 6-8)

- Knows that energy comes in different forms, such as light, heat, chemical, nuclear,
- mechanical and electrical
- Understands that energy cannot be created or destroyed, but only changed from one form to
- another
- Knows that the sun is a major source of energy for changes on the earth's surface; the sun's
- energy arrives as light with a range of wavelengths consisting mainly of visible light with
- significant amounts of infrared and ultraviolet radiation
- Knows that heat energy moves in predictable ways, flowing from warmer objects to cooler
- ones until both objects are at the same temperature
- Knows that heat can be transferred through materials by the collision of atoms or across
- space by radiation; if the material is fluid, currents will be set up in it that aid the transfer of
- heat
- Knows that electrical circuits provide a means of converting electrical energy into heat, light,
- sound, chemical, or other forms of energy
- Knows that in most chemical reactions, energy is released or added to the system in the form
- of heat, light, electrical, or mechanical energy

Level IV (Grades 9-12)

- Knows that although energy can be transferred by collisions or waves and converted from
- one form to another, it can never be created or destroyed, so the total energy of the universe
- is constant
- Knows that all energy can be considered to be either kinetic energy (energy of motion),

- potential energy (depends on relative position), or energy contained by a field
- (electromagnetic waves)
- Knows that heat energy consists of random motion and the vibrations of atoms, molecules,
- and ions; the higher the temperature, the greater the atomic or molecular motion
- Knows that energy tends to move spontaneously from hotter to cooler objects by
- conduction, convection, or radiation; similarly, any ordered state tends to spontaneously
- become less ordered over time
- Knows that the energy of waves (electromagnetic and material) can be changed into other
- forms of energy (e.g., chemical and electrical), just as other forms of energy (chemical and
- nuclear) can be transformed into wave energy
- Knows that some changes of atomic or molecular configuration require an input of energy,
- whereas others release energy
- Knows that each kind of atom or molecule can gain or lose energy only in particular discrete
- amounts and thus can absorb and emit light only at wavelengths corresponding to these
- amounts; these wavelengths can be used to identify the substance
- Knows that fission is the splitting of a large nucleus into smaller pieces, and fusion is the
- joining of two nuclei at extremely high temperature and pressure; nuclear reactions convert a
- fraction of the mass of interacting particles into energy

A school or district that takes the "exemplar" approach to benchmarks would require students to demonstrate competence in a selected number of benchmarks per level. For example, a school or district might require students to demonstrate competence in two out of the three benchmarks for Level I; three out of five for Level II; five out of seven for Level III; and six out of eight for Level IV. This approach would allow a school or district to meet a larger number of standards without exceeding the recommended limit of 600 benchmarks discussed in the preceding section. It would also allow for more flexibility within the classroom, in that individual teachers would have the option to use those benchmark components that they judged most applicable for their students. However, this approach also results in less continuity of coverage within a content domain since different teachers will no doubt select different benchmark exemplars to illustrate student competence within the levels for a given standard. It is also important to note that this approach may defeat the designed purposes of some well-articulated standards, such as those developed by Project 2061, where upper-level benchmarks are predicated under the assumption that students are familiar with a logically prior concept addressed at an earlier level. If teachers select without regard to articulation, some of the value of this approach may be lost.

3. Will student performance be reported using course grade or standards?

Currently, most schools and districts report student progress using appropriate grades for broad academic areas organized within courses. However, current research and theory indicate that courses of the same title do not necessarily cover the same content. (Yoon, Burstein & Gold, not dated) In other words, two courses of the same name do not necessarily cover the same declarative, procedural, and contextual knowledge. If a school or district wished to use traditional grades but implement a standards-oriented approach, it would ensure that the benchmarks that have been identified would be distributed systematically throughout the various courses within content areas--that is, specific benchmarks would be assigned to courses based upon the elements they cover. Any two courses with the same title would not only cover the same benchmarks, but would place the same relative importance on the benchmarks they cover.

For example, assume that two courses of the same title were designed to cover the same seven benchmarks. The school or district could also determine which percentage of the grade each benchmark would command. In such a case, it might be determined that the first two benchmarks each accounted for 25% of the grade and the remaining five benchmarks accounted for 50% of the grade. Clearly, this would provide more precision for course descriptions and show an equivalence between "identical" courses that is not often found today.

In short, traditional grading practices and standards-based assessment are not incompatible. A school or district must simply distribute and weight the standards that have been identified across the various courses in a systematic, well-reasoned fashion.

The second reporting option a school or district might adopt is to report student progress by benchmarks. Rather than assign a single grade to a course, a teacher would report progress in some way for each benchmark covered in the course. In effect, for assessment purposes only, each benchmark component would be considered independent of the others covered within the course. When this approach is taken, schools and districts commonly employ rubrics as opposed to grades. A rubric is a description of the levels of understanding or skill for a given benchmark. For example, below is a rubric for the Level II mathematics benchmark "*Understands the basic role of place value .:*"

4. Demonstrates a thorough understanding of the role and function of place value and provides insights that are not obvious when using the concept of place value.
3. Demonstrates a complete and accurate understanding of the role and function of place value as it relates to estimating or calculating addition, subtraction, multiplication, and division.
2. Displays an incomplete understanding of the role and function of place value as it relates to estimating or calculating addition, subtraction, multiplication, or division.
1. Has severe misconceptions about the role and function of place value as evidenced by severe place value errors in addition, subtraction, multiplication, or division.

Commonly, one of the described levels within a rubric is designated as the targeted level of skill or knowledge. For example, a score of 3 in the reporting rubric above might be selected as the target standard for the Level II mathematics benchmark, "*Understands the basic role of place value.*"

Reporting out by benchmarks would, of course, require a record-keeping system that is far different from that currently used in most schools and districts. Each student's score on individual benchmarks would be recorded. Assuming the use of a four-point rubric, individual students would receive a score of 1 through 4 on each benchmark assessed within each standard. These scores could then be averaged to obtain an overall standard score at a given benchmark level.

4. Will all students be required to meet all standards?

A major decision facing a state, school or district that wishes to emphasize content area standards is whether students will be required to meet a targeted level of knowledge and skills. This approach is reminiscent of the mastery learning approach of the 1970s and early 1980s (see Levine & Associates, 1985) and the more recent outcomes-based approach, or OBE approach (Spady, 1988). In the context of the reporting rubric described previously, a mastery or outcomes-based approach would mean that students would be required to receive a score of 3 on each benchmark. If a student did not meet the targeted level for a benchmark (i.e., did not obtain a score of 3 on the rubric), he or she would be provided with additional instructional opportunities until he or she could meet the required proficiency. Of course, such a system can make extreme demands on resources. In a traditional system, no extra resources need be used if a student does poorly in a course. In a mastery or OBE system, each student who does not meet a standard is provided with whatever instructional and curriculum resources are necessary to ensure that the student meets the requirements. A variation in the theme of a comprehensive mastery or outcomes-based approach is to require that students meet the performance standards on some, but not all, benchmarks. Those benchmarks that are applied to all students would be considered a set of core requirements.

Conclusion and References

In this monograph, we have attempted to build a rationale for organizing educational reform around standards. Even though the standards movement at the national level has bogged down and state and district level efforts to effect such reform have been extremely uneven, standards-based education has a logic that is to us compelling. However, as compelling as that logic is, the process of designing a standards-based educational system is a difficult and technical task. We have described a data base, developed at the Mid-continent Regional Educational Laboratory, that we believe alleviates much of the technical detail work. We have also described a set of questions that, among many others, must be addressed if a standards-based approach is to be effectively implemented. We hope that our efforts will facilitate the transformation of American education to a system that holds itself and students accountable to specific standards of knowledge and skill.

References

American Association for the Advancement of Science (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Bennett, W. J. (1992). *The devaluing of America: The fight for our culture and our children*. New York: Summit Books.

Berliner, D. C. (1979). Tempus Educare. In P. L. Peterson & H. J. Walberg (Eds.), *Research on teaching*. Berkeley, CA: McCutchan.

Berliner, D. C. (1984b). The half full glass: A review of research in teaching. In P. L. Hosford (Ed.), *Using what we know about teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Brandt, R. (1995). Overview: What to do with those new standards. *Educational Leadership*, 52(6), p. 5.

Cohen, D. (1995). What standards for national standards? *Phi Delta Kappan*, 76(10), 751-757.

Coleman, J. S. (1972). The evaluation of equality of educational opportunity. In F. Mosteller & D. P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 140-161). New York: Vintage Books.

DiegmueLLer, K. (1995, April 12). Running out of steam. *Struggling for standards: An Education Week special report*. Washington DC: Education Week.

Dorr-Bremme, D. W., & Herman, J. L. (1986). *Assessing student achievement: A profile of classroom practices*. Los Angeles: Center for the Study of Evaluation, University of California, Los Angeles.

Durm, M. W. (1993). An A is not an A is not an A: A history of grading. *The Educational Forum*, 57(Spring), 294-297.

Eisner, E. W. (1995). Standards for American schools: Help or hindrance? *Phi Delta Kappan*, 76(10), 758-764.

Finn, C. E., Jr. (1990). The biggest reform of all. *Phi Delta Kappan*, 71(8), 584-592.

Fisher, C. W., Filby, N., Marliave, R. S., Cahen, L. S., Dishaw, M. M., Moore, J. E., & Berliner, D. C. (1978). *Teaching behaviors, academic learning time and student achievement*. Final report of phase III-B. Beginning Teacher Evaluation Study. San Francisco, CA: Far West Laboratory of Educational Research and Development.

Gandal, M. (1995). *Making standards matter: A fifty-state progress report on efforts to raise academic standards*. Washington, DC: American Federation of Teachers.

Glaser, R. & Linn, R. (1993). Forward. In L. Shepard, *Setting performance standards for student achievement* (pp. xiii-xiv). Stanford, CA: National Academy of Education, Stanford University.

- Guskey, T. R. (Ed.). (1996). *Communicating student learning: 1996 ASCD yearbook*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Harnischfeger, A. & Wiley, D. (1978). *Conceptual and policy issues in elementary school teaching: Learning*. Paper presented at the Annual Meeting of the American Educational Research Association, Toronto.
- Kendall, J. S. & Marzano, R. J. (1996). *Content knowledge: A compendium of standards and benchmarks for K12 educators*. Aurora, CO: Mid-continent Research for Education and Learning.
- Levine, D. V. & Associates (1985). *Improving student achievement through mastery learning programs*. San Francisco, CA: Jossey-Bass.
- Lewis, A. C. (1995). An overview of the standards movement. *Phi Delta Kappan*, 76(10), 744-750.
- Mager, R. (1962). *Preparing instructional objectives*. Palo Alto, CA: Fearon Publishers.
- Marzano, R. J. & Kendall, J. S. (in press). *How to design a standards-based district, school, or classroom*. Aurora, CO: Mid-continent Research for Education and Learning.
- National Association of State Boards of Education (1988). *Rethinking Curriculum*. Alexandria, VA: Author.
- National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Department of Education.
- National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Education Commission on Time and Learning (1994). *Prisoners of time*. Washington, DC: Author.
- National Education Goals Panel (1991). *The national education goals report: Building a nation of learners*. Washington, DC: Author.
- National Education Standards and Improvement Council (1993). *Promises to keep: Creating high standards for American students*. A report on the review of education standards from the Goals 3 and 4 Technical Planning Group to the National Educational Goals Panel. Washington, DC: National Goals Panel.
- National History Standards Project (March 1993). *Progress report and sample standards*. Los Angeles, CA: National Center for History in the Schools.
- National Update on America's Education Reform Efforts (1995, November 20). *The National Report Card* (Vol. 5, No. 62).
- Nixon, R. M. (1970). *Special message to the Congress on educational reform* (March 3). Washington, DC.

- Olson, L. (1995). Cards on the table. *Education Week (June 14)* (pp. 23-28).
- O'Neil, J. (1995). On using standards: A conversation with Ramsay Seldon. *Educational Leadership*, 52(6), 12-14.
- Popham, W. J. (1972, April). Must all objectives be behavioral objectives? *Educational Leadership* (pp. 605608).
- Popham, W. J. (1994). The instructional consequences of criterion-referenced clarity. *Educational Measurement: Issues and Practices*, 13(4), 15-18;30.
- Ravitch, D. (1983). *The troubled crusade: American education 1945-1980*. New York: Basic Books.
- Ravitch, D. (1995). *National standards in American education: A citizen's guide*. Washington, DC: Brookings Institute.
- Resnick, L. & Nolan, K. (1995). Where in the world are world-class standards? *Educational Leadership*, 52(6), 6-10.
- Robinson, G. E. & Craver, J. M. (1989). *Assessing and grading student achievement*. Arlington, VA: Educational Research Service.
- Shepard, L. (1993). *Setting performance standards for student achievement*. Stanford, CA: National Academy of Education, Stanford University.
- Spady, W. G. (1988). Organizing for results: The basis of authentic restructuring and reform. *Educational Leadership*, 46(2), pp. 4-8.
- Tucker, M. (1992, June 17). A new social compact for mastery in education. *Education Week: Special Report* (p. S3).
- Tyack, T. & Tobin, W. (1994). The "Grammar" of schooling: Why has it been so hard to change. *American Educational Research Journal*, 31(3), 453-479.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. University of Chicago Press.
- Tyler, R. W. (1932, 1989). Constructing achievement tests. In G. F. Madaus & D. Stefflebeam (Eds.), *Educational evaluation: Classic works of Ralph W. Tyler* (pp. 17-86). Boston: Kluwer Academic Publishers.
- Wiggins, G. (1993, November). Assessment, authenticity, context and validity. *Phi Delta Kappa*. (pp. 200-214).
- Yoon, B., Burstein, L., & Gold, K. (undated). *Assessing the content validity of teacher's reports of content coverage and its relationship to student achievement*. (CSE Report No. 328). Los Angeles, CA: Center for Research in Evaluating Standards and Student Testing, University of California, Los Angeles.

September 26, 1997