

Occasional Paper No. 1

Standards for Science Education

Senta A. Raizen



Funded by the
National Science Foundation

National Institute for Science Education (NISE) Publications

The NISE issues papers to facilitate the exchange of ideas among the research and development community in science, mathematics, engineering, and technology (SMET) education and leading reformers of SMET education as found in schools, universities, and professional organizations across the country. The NISE Occasional Papers provide comment and analysis on current issues in SMET education including SMET innovations and practices. The papers in the NISE Research Monograph series report findings of original research. The NISE Conference and Workshop Reports result from conferences, forums, and workshops sponsored by the NISE. In addition to these three publication series, the NISE publishes Briefs on a variety of SMET issues.

The research reported in this paper was supported by a cooperative agreement between the National Science Foundation and the University of Wisconsin–Madison (Cooperative Agreement No. RED-9452971). At UW–Madison, the National Institute for Science Education is housed in the Wisconsin Center for Education Research and is a collaborative effort of the College of Agricultural and Life Sciences, the School of Education, the College of Engineering, and the College of Letters and Science. The collaborative effort is also joined by the National Center for Improving Science Education, Washington, DC. Any opinions, findings, or conclusions are those of the author and do not necessarily reflect the view of the supporting agencies.

Occasional Paper No. 1

Standards for Science Education

Senta A. Raizen

National Institute for Science Education
University of Wisconsin-Madison

July 1997

About the Author

Senta A. Raizen is Director of the National Center for Improving Science Education in Washington, DC. She has written numerous books and reports on science education in elementary, middle, and high school, as well as on preservice education of elementary school teachers and technology education. She serves as an advisor to the National Assessment of Educational Progress, the National Goals panel, and the Third International Mathematics and Science Study.

Contents

Abstract	v
Why Science Education Standards?.....	1
Changed Goals	1
Reform Strategies.....	2
The NCTM Standards	3
What Are Standards?.....	5
Engineering versus Ecological Metaphors.....	6
Development of the Science Standards	7
New Views of Science Learners	8
The Proactive Organizations	8
The Science Content of <i>NSES</i> , the <i>Benchmarks</i> , and <i>The Content Core</i>	11
A Content Comparison	12
Excerpts from <i>Benchmarks for Science Literacy</i>	15
Excerpts from <i>NSES</i>	17
Excerpts from <i>The Content Core</i>	19
Concordance between <i>Benchmarks</i> and <i>NSES</i>	22
Supporting Standards in <i>NSES</i>	23
Excerpts from Science Teaching Standards	23
Excerpts from Assessment in Science Education Standards	24
Excerpts from Science Education System Standards	25
Assessment Standards	26
The New Standards Project	26
The National Assessment of Educational Progress (NAEP)	31
The Third International Mathematics and Science Study (TIMMS)	34
Analysis of University Entrance Examinations	35
Science Standards in the States	37
The Current Status of State Science Curriculum Frameworks	37
Variation Among Frameworks	38
Frameworks, Standards, and Reform	40
References	41

Abstract

After providing the current rationale and historical background for educational standards, the paper discusses various meanings and interpretations attached to this term. It then provides a comparative analysis of three sets of publications that are seen as providing national standards for science education, developed by the National Science Teachers Association, the American Association for the Advancement of Science, and the National Research Council. Next, the role of assessment in setting standards is discussed, in particular, the science frameworks used by the National Assessment of Educational Progress and the Third International Mathematics and Science Study, the work of the New Standards Project, and the expectations built into rigorous university entrance exams, as represented by the Advanced Placement examinations. The paper also addresses the current status of state science curriculum frameworks, including commonalities and variations among them. It concludes with a brief discussion of standards and framework documents, whether nationally or state developed, as policy levers for reforming science education in elementary and secondary school.

This paper addresses the development, content, and potential for implementation of nationally developed science education standards, including their adoption and effects at the state level. Specifically, the focus will be on:

- the *National Science Education Standards* published in early 1996 by the National Research Council (NRC),
- the two landmark documents—*Science for All Americans* (SFAA) and *Benchmarks for Science Literacy*—published by the American Association for the Advancement of Science's (AAAS) Project 2061 in 1989 and 1993,
- *The Content Core* document published by the National Science Teachers Association's (NSTA, 1992) Scope, Sequence, and Coordination (SS&C) Project, and
- the performance assessment materials developed by the New Standards Project (NCEE, 1995 a, b, c).

I will also touch on such other standards-related activities as the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study (TIMSS), an analysis of university entrance examinations in science given in other countries and the U.S., and the status of state curriculum guidelines in science. Before taking up these topics, however, it is necessary to set the creation of science education standards in the wider context of the last decade's education reform and policy.

Why Science Education Standards?

The drive for education standards appears to have several roots, among them current educational goals in the face of a changing student body, current conceptions of reform that emphasize systemic approaches and accountability mechanisms, and the fortuitous appearance of the *Curriculum and Evaluation Standards for School Mathematics* developed by the National Council of Teachers of Mathematics (NCTM, 1989), which provided a much-needed existence proof that the creation of national standards was possible in the U.S.

Changed Goals

The emphasis in current reform efforts in science (and mathematics) education is on science (mathematics) for all, as exemplified by the titles of key reform publications: AAAS's *Science for All Americans* and *Benchmarks for Science Literacy*; the NRC's (1989) *Everybody Counts*, associated with the mathematics standards developed by NCTM; the "Call to Action" section of NRC's *National Science Education Standards*, which proclaims the purpose of that document to be making scientific literacy for all a reality; and the implementation of SS&C, which has as its motto (at least in California): Every student, Every science, Every year. This is considerably different from the impetus for the 1960s reforms, occasioned by the launching of Sputnik by the USSR in October 1957, which appeared to pose a technological and military challenge to the U.S. The response then was the expansion of a scientific cadre capable of overcoming and surpassing the Russian achievement, hence the emphasis on providing up-to-date curricula and well-trained teachers in order to ready a greater number of students for majoring in the sciences and entering scientific careers (Raizen, 1991).

The current more inclusive goal responds to the perceived need for (a) the U.S. to remain competitive in the world economy, which is seen as requiring a workforce in which everyone—not just Ph.D. scientists—is adequately educated in science and mathematics; (b) the average citizen to understand enough science to deal in an informed way with individual, family, and community decisions (e.g., concerning health or environmental issues); and (c) ensuring access to scientific careers for all students so motivated, no matter their gender, ethnicity, or socioeconomic status.

The goal of scientific literacy for all will not be easy to achieve. Although this view is not unanimous (see, for example, Bracey, 1993), today's U.S. students generally are perceived to be doing poorly in international assessments that compare their mathematics and science achievement to those of students in other industrialized nations (IEA, 1988; Lapointe, Mead, & Phillips, 1989, 1992), or even to earlier student achievement within the U.S. (Mullis, Dossey, Foertsch, Jones, & Gentile, 1991).

Moreover, the composition of the student body is changing—increasing in just those population groups that have not fared well under traditional science education methods. In California, for example, minority students constituted 27% of the 1970 school population. By 1980 the proportion of minority students had risen to 42%, and by 2000, minorities in California public schools are expected to surpass Anglos, comprising 52% of the school population (Catlin, 1986). While California may be setting the pace, the rest of the nation is not far behind. The percentage of Black and Hispanic students enrolled in grades 1-12 in central city public schools increased from 42% to 52% during the two decades between 1970 and 1989 and from 10.6% to 20% in other metropolitan schools (National Center for Education Statistics, 1992). Added to these increases of students from minority groups are increases of students whose native language is not English: 23% of the 5- to 17-year-olds in California public schools speak a language at home other than English; the percentage is nearly the same in suburban jurisdictions adjoining large cities (e.g., in Arlington County, Virginia, where 3,700 students out of 17,500 come from non-English speaking families). Nationally, between 1980 and 1990, students who spoke a language other than English at home increased from 4.5 million to 6.3 million or from 10% to 14% of all children (NSF, 1996, p.5). More than 40 different languages were represented, ranging from various forms of Spanish and Chinese to Tagalog and Hmong.

Thus, high standards for all students are seen both as a response to national and personal needs and as a way to address equity issues reflected in the low proportion of Blacks, Hispanics, and students from impoverished families and communities represented in the scientific and technological workforce.

Reform Strategies

Two "lessons" from the 1960s reform efforts in science education reinforced the belief in nationally developed standards as a critical reform tool. The first lesson concerned the seeming ineffectiveness of focusing on specific improvement strategies, for example, teacher institutes, curriculum improvement, or student enrichment. Experience over several decades of effort indicated the need to build a common vision of what made for good science education and then coordinate activities across potential intervention strategies around that vision instead of

pursuing various of them in isolation.¹ This kind of thinking gave rise to the notion of systemic reform (Smith & O'Day, 1991; O'Day & Smith, 1993), based on the belief that changing any one component of an education system is insufficient to cause significant reform. Instead a concerted effort to change the structure, operating procedures, forms of interaction, and the distribution of power is needed before any real change can transpire in a system. The objective is to make a comprehensive set of changes work together, targeting the education system at the state or district level, even though change still has to happen classroom by classroom. For example, NSF guidelines for systemic initiatives call for alignment of policy and practice with demanding goals for student learning, and with particular attention to developing competent teachers (through both inservice and preservice programs), challenging curriculum content and up-to-date teaching practice, assessment of student learning in line with the goals and curricular content, and appropriate uses of educational technology. Active partnerships with stakeholders inside and outside education are seen as critical to systemic improvement efforts. And an integral part of systemic reform is the formulation of a vision for education, generally in the form of standards.

A second lesson from the 1960s was the difficulty of instituting improvements over the wide range of schools and school districts in the United States, each locally governed. Results of international comparisons of student achievement in science and mathematics seemed to underscore the effectiveness of centrally controlled education systems, such as those of Japan and France, contrasted to the multiple, complex, and often incoherent levels of governance characterizing U.S. education that frequently emit conflicting policy signals. Notable is the strong influence in some centralized systems of national curriculum standards that lay out what students are to learn at given grade levels. The apparent consensus of the early 1990s that the U.S., too, needs national standards (standards envy?) seems to be eroding somewhat at present (Ravitch, 1995). The spring 1996 Governors' conference, a replay of the 1990 conference that established national education goals (NGA, 1990), reaffirmed the localism that characterizes U.S. education by assigning the responsibility for formulating standards to the states (rather than any national bodies), the intent being that such standards would provide criteria for local school personnel to decide on the specifics of what should be taught and how it should be taught. Presumably, curriculum and instruction need not be the same for all children or sites in order to implement the goal of high standards for all. By the same token, a number of states have instituted state assessments to gauge the extent to which student learning goals are being achieved.

The NCTM Standards

A potent catalyst for the efforts to develop standards was provided by the National Council of Teachers of Mathematics (NCTM) in their curriculum standards project, the results of which were published in 1989. (For a detailed history of this effort, see the case study report by McLeod et al. in Raizen & Britton, 1996). The release of the first standards document, the *Curriculum and Evaluation Standards* (NCTM, 1989), was accompanied by a publicity

¹ The perception of the innovations of the 1960s as a collection of isolated projects is not quite accurate. Thus, in the 1970s, NSF began to stress integration of teacher training with curricula; Welch (1979) notes that, by 1975, 80% of funding for training teachers was devoted to their learning to implement new curricula.

campaign so successful that it astonished even the officers of the organization as well as the writers who had contributed to the work, but perhaps the main reason for NCTM's success was timing. As Ravitch (1995) points out, "At the very time that governors and other political leaders wondered about the feasibility of voluntary national standards, there were the NCTM Standards as an example for emulation" (p. 57).

Documents that set curricular goals in specific school subjects are not really a new development in education, except at the national level. A number of states have had curriculum guides in place for many years; in fact, there is a reasonably close relationship between California's 1985 mathematics framework (California Department of Education, 1985) and the NCTM *Curriculum and Evaluation Standards*, the first and best-known of the NCTM standards documents. Two additional ones have been published subsequently: *Professional Standards for Teaching Mathematics* (NCTM, 1991) and *Assessment Standards* (NCTM, 1995). To quote from the NCTM case study:

Given their task, writers naturally looked at other statements of curricular goals, including curriculum guidelines from California, Oregon, Wisconsin, and other states. The California *Framework* was mentioned frequently. As one writer noted: "Certainly the 1985 California *Framework* was one of the documents that was used in helping to formulate the NCTM *Standards*. It was something that everybody in all of the groups was familiar with and looked at for help in thinking about what the *Standards* might contain."

A leader from California noted that he was careful not to push too hard since people will reject an idea "just because it comes from California." But he did think that California was a major influence:² ". . . there was a very big influence from the mathematics education community in California. . . . When I look back, I think that the 1985 California *Framework* really set the stage for a lot of things." (McLeod et al., 1996, p. 39).

Why did the NCTM standards effort, and particularly the first document on curriculum and evaluation, become so widely known and accepted? One reason was its style—it set out a vision of mathematics in the schools that was understandable to the field, yet not too detailed. As a state supervisor of mathematics interviewed and quoted in the case study said:

One of the brilliant characteristics of the *Curriculum Standards* is that the grain size is big enough that the bullets aren't damaging; they cannot be treated like behavioral objectives. Brilliant move! They're not sweepingly general; they can't be rejected because they're too general. But they're not at the small grain size where you'd have 150 per grade span, [the grain size] that people are used to. (McLeod et al., 1996, p. 116)

Another reason is the long and careful developmental period that eventually led to the formulation of the 1989 *Standards*. An earlier policy document, NCTM's (1980) *An Agenda for Action*, used during the early 1980s had been the focus of NCTM's national and regional

² Questionnaire data confirmed this view; the writers rated the California *Framework* just behind *An Agenda for Action* (NCTM, 1980) in its influence on the *Curriculum and Evaluation Standards*.

meetings and professional development activities. Also, the development of all the standards documents was a very inclusive process. Teachers, mathematics supervisors, teacher education faculty, and others with expertise in research, technology, or other areas made up the writing teams (Crosswhite, Dossey, & Frye, 1989). This ensured that realism based on classroom experience and knowledge of earlier reform efforts would temper the document's visionary recommendations, making for wide acceptance in the field.

NCTM also was able to gain support from all the major professional associations in mathematics (Crosswhite, Dossey, & Frye, 1989). In that respect, mathematics may be in a more favorable position than other school subjects, with relative unanimity within the community of mathematicians and mathematics educators on the content and sequence of mathematics instruction from kindergarten through high school and beyond—hardly the case in the sciences or other school subjects. The difficulty that other groups have had in developing standards (Donmoyer, 1995; Myers, 1994) only serves to amplify this distinctive characteristic of mathematics as a school subject.

Without attention to the long period of development enjoyed by the NCTM standards, the special characteristics of mathematics as a field, and the great publicity that surrounded the release of the 1989 *Standards*, the federal government seized on their success to sponsor development of standards in all the key school subjects, including the sciences, history, English language, social studies, reading, and the arts.

What Are Standards?

The word *standards* encompasses several meanings; possibly this ambiguity makes it both more acceptable and more powerful. Although within NCTM the standards work originated from concerns with the claims made by some mathematics textbook authors, the term came to be used in the sense of a banner, a rallying flag for professionals in mathematics education. To some extent, this view of the purposes of a standards document was in reaction to the back-to-basics movement and the spread of minimum competency testing that characterized the 1970s, when behaviorist psychology began to influence educational philosophy. Curriculum was stated in fine-grained behavioral objectives, based on the work of R. M. Gagné (1970) in mathematics and science education. The emphasis on testing for basic skills further reinforced the specification of curriculum guidelines in terms of measurable objectives. States and districts, as well as curriculum developers and textbook authors, were expected to specify these objectives in considerable detail. In contrast, the NCTM *Standards* were to exemplify contemporary thinking about best practice in mathematics education, intended as much to inspire as to prescribe. NCTM further defines standards as serving three purposes: minimal criteria for quality, an expression of expectations of goals, and means for leading a group toward new goals (NCTM, 1989). A similar approach was taken by the California mathematics and science frameworks (California Department of Education, 1985, 1990). Bill Honig, at the time the state's chief state school officer, said (as quoted in Atkin, Helms, Rosick, & Siner, 1996, p. 21): "The [science] framework was to give focus [to the California science education reform] without giving explicit direction."

For many of those shaping education policy, however, the notion of standards implies a mechanism by which to hold schools accountable for what students learn. They see standards as

a set of prescriptions for schools, teachers, and students. As Atkin (1994, p. 82) put it: "Standards' has the kind of bite politicians like." The accountability emphasis is very much in the current spirit of school reforms, and this interpretation certainly has made for ready acceptance of the notion of standards among federal and state policymakers, including the nation's governors who had articulated the national education goals.

These two different interpretations lead to quite different types of documents, however, especially in the level of detail—"grain size," to borrow the mathematics supervisor's expression. The "vision" or "banner" interpretation allows for a document vague enough to collect many supporters, yet specific enough to offer something to practitioners (Apple, 1992). As Donmoyer (1995) points out, using a large grain size for standards documents may make for acceptability by the various interest groups, but it leaves all the hard decisions for later. The need to provide adequate guidance for teachers has led the American Federation of Teachers, a strong advocate of rigorous standards, to list sufficient specificity as one of its criteria for effective state standards to ensure "the development of a common core curriculum" (AFT, 1995), including specification by grade or age bands.

Many teachers, particularly at the elementary level, may need quite a bit of precision with respect to curriculum content in a state framework. A solution adopted by California, which had led the nation and the states in developing innovative mathematics and science state frameworks, is for each school to develop a content matrix to create an articulated, coherent science program. This is in the spirit of combining national or state standards with local options in implementation and banks on teachers being professionals with the capacity to define both the content and the instructional approaches for the subjects they teach. However, it is likely to exacerbate teachers' workloads and raise concerns about the fidelity and, hence, effectiveness of the reform effort. There are, for example, some contradictory findings about the implementation of the California mathematics framework, which seemed to be going well according to studies reporting on the spread of knowledge and engagement of teachers around the framework (Findings, 1990), but not going so well with respect to actual translation into improved classroom practice (Marsh & Odden, 1991).

If standards are to serve purposes of accountability, they must be sufficiently detailed to allow derivation of performance standards for student achievement and development of appropriate assessments to measure that achievement. And in response to the accountability demands, various stakeholders have been calling for yet additional standards, particularly opportunity-to-learn standards, teacher development standards, and program standards. These demands are premised on the view that schools and teachers must be provided the wherewithal to enable students to achieve the desired performance standards. In short, ambiguity about the term *standards* persists and already has led to controversy when the different meanings begin to clash.

Engineering versus Ecological Metaphors

From physics and engineering through the life sciences and social sciences, scientists conceive of the phenomena they are studying as embedded in systems. However, the nature of these systems and our ability to simplify them, tease apart complex interrelationships among various components, and quantify these relationships vary greatly from field to field. Interestingly, the language that accompanies the current systemic reform efforts in science education and, specifically, the formulation and application of standards is that of precisely engineered systems rather than language appropriate to ecological systems that evolve over time. Project 2061, one of the three key organizations involved in the formulation of national science standards, includes in its reform strategies the development of "designs," "blueprints," and "benchmarks." In the language associated with the main federal sponsors of national and state standards (NSF and the Department of Education), curriculum standards are seen to be the "drivers" of reform; tests and textbooks are to be "aligned" with the standards, as are teacher preservice and inservice education (NSF, 1995). State frameworks based on the national standards are to be created, disseminated, and implemented and, thus, to improve student outcomes. These outcomes are to be clearly defined through performance standards for students as well as for schools and districts, on the basis of which accountability systems would be established to monitor performance. Both the conception and the terms are borrowed from quality controls applied in systems engineered to provide precisely specified goods or services. This mechanistic model of designed change assumes a tightly coupled system in which the causal links between individual parts are well understood—surely not the case for education systems. The metaphors used in reform efforts are not without consequence. As Eisner (1992) says:

When the language of industrial competition is used to make a case for particular educational aims—"losing our competitive edge"—our conception of the mission of schools is gradually shaped in industrial terms. The school becomes viewed as an organization that turns out a product—a student—whose features are subject to the same quality control criteria that are applied to other industrial products. (p. 303)

One is reminded of an earlier era that saw the introduction of Taylorism into American education around the turn of the century (Callahan, 1962). Then, too, schools were not living up to the new demands placed on them, and one of the reasons was that they were perceived to be inefficient. Taylor's industrial paradigm called for defining standardized goals for schooling, so that student outcomes (the "products") could be appraised accordingly and the performance of teachers (the "workers") managed and judged to achieve the desired product quality. Interestingly, the very nature of what is being called for in science education reform is antithetical to such standardization, as was Dewey's formulation of progressive education, the purpose of which was the cultivation of unique talents.

Development of the Science Standards

As with the NCTM mathematics standards, the reform efforts that eventually led to the development of science standards were originated by science educators, rather than by more generalized policy initiatives in education. The reasons for a new round of science education reforms starting in the 1980s were the seeming disaffection of many students with science

instruction, leading to poor achievement as well as low enrollments, and accumulating evidence on how science is learned, with the attendant promise of being able to improve on current instructional practices. Contributing also to the perceived need for reform were that the sciences themselves were changing and that there were new opportunities offered by the developing educational technology.

New Views of Science Learners

Two different streams of research have combined to make up current conceptions of science learning. The first comes from a body of cognitive research, by now quite considerable (e.g., Driver et al., 1985; Harlen, 1985, 1992; Shymansky et al., 1990; Tobin et al., 1988), which has provided evidence that learners are not passive recipients of codified knowledge but rather active participants who bring their own preformulated notions of the natural world to the science classroom. Second, current epistemological views of science hold that the development of scientific knowledge occurs through consensus building within professional communities that agree on methodology, the nature of evidence and proof, and the kind of discourse appropriate to their field.

At the same time, both the problems that scientists address and the ways in which they are able to address them have changed profoundly, in part because of the advent of powerful computers and other information technology and in part because of societal demands as expressed through the availability of public funding for research. Research on problems of visible societal import are more likely to be funded; generally such research cuts across several subspecialties and involves large teams across national borders, as exemplified by the Human Genome project that is to identify the many millions of human genes. The changing nature of science has served to reinforce the keystones of science education reform, which can be discerned as well in the development of the standards. Essentially, the reform initiatives emphasize one or more of the following:

1. instruction through presentation of real-world problems and applications rather than abstract knowledge;
2. providing opportunities for students to investigate natural phenomena, often involving the use of computers to manage data; and
3. explicating linkages across fields of science and to other subjects, including mathematics.

Other strands common to education reform in general can be seen in the science reforms and standards as well, including learning to work in groups and to communicate effectively—both in spoken and written natural language and in symbolic form.

The Proactive Organizations

Three organizations have been key to the development of science education standards: the National Science Teachers Association (NSTA), the American Association for the Advancement of Science (AAAS) through its Project 2061, and the National Research Council (NRC). Though NRC ultimately accepted the responsibility (and the funding) for developing the standards, the other two organizations played major roles. It is perhaps noteworthy that two of these

organizations are bodies essentially constituted of scientists, whereas the third encompasses largely teachers of science and science educators.

AAAS. Project 2061 was the first to publish a widely disseminated vision statement: *Science for All Americans (SFAA)*. This document was written by the staff of Project 2061 based on the reports of five panels, one each in the biological and health sciences, physical and information sciences and engineering, mathematics, social and behavioral sciences, and technology. While panel members were largely drawn from academic research, some scientists and engineers from industry also participated. *SFAA* was widely reviewed, with about half the reviewers comprising teachers, science education specialists, and nonresearch scientists.

Having set out the vision of what all Americans should know in science, Project 2061 sought to avoid the dilemma of inappropriate or ineffective interpretation at the classroom level. Working with six school district-based centers, it counted on teachers and teacher leaders to translate *SFAA* into curricula that could be implemented in the classroom. Though the initial notion was that the six centers would develop alternative curriculum models for the rest of the country to adopt and adapt, the Project 2061 central staff found that an intermediate step was necessary, namely "backmapping" from *SFAA* to define what students would need to know at various grade levels and in what sequences this knowledge should be taught in order for them to emerge as the scientifically literate persons envisaged in that document. This process proved to be difficult, but eventually each team produced maps for many of the learning outcomes, with considerable agreement among the sites, according to the project staff (AAAS, 1992).

Although these maps originally were intended as an internal product to advance the curriculum work of the six centers, they became the basis for Project 2061's second landmark document, *Benchmarks for Science Literacy* (AAAS, 1993). This document was widely disseminated and regarded by many as the science standards analogous to the NCTM mathematics standards, despite the fact that the NRC had been funded to develop *the* national standards. A recent examination of newly developed state science frameworks (SRI International, 1996) revealed that a number of them refer to the *Benchmarks* as a cornerstone document in their own formulation of standards. In fact, one state went so far as to adopt the whole *Benchmarks* document outright, fleshing out a few selected portions and then inviting the teachers in the state to follow the model provided in the state guidelines to detail the rest of the science curriculum based on the *Benchmarks*.

NSTA. The executive director of the NSTA proposed a plan in 1989 for secondary science education that has become known as Scope, Sequence, and Coordination, or SS&C (Aldridge, 1989). The plan called for (a) science for all students every year of grades 7-12 (later expanded to include grade 6 as well); (b) coordination across biology, chemistry, earth/space science, and physics so that students would be able to see the interrelationships and applications of important concepts; and (c) "spaced learning," revisiting concepts on a periodic basis so as to treat them with increasing depth and sophistication.

Having proposed such a radical reform of the current structure of secondary school science, NSTA found it necessary to provide greater detail on how the required restructuring might be accomplished. *The Content Core* (NSTA, 1992) was developed for curriculum designers with two major purposes: to outline the science content to be taught over the seven recommended

grades, and to illustrate methods for coordinating and integrating the science curriculum across the four fields of science. In addition, some strategies for implementation are provided in the monograph. As with any standards document, *The Content Core* is not a curriculum per se, but is intended "as a template for designing courses, selecting instructional materials, and constructing assessment instruments" (p. 7).

The foreword to the document points out its compatibility with AAAS's Project 2061 and the tenets proclaimed in *SFAA*, particularly its theme of "less is more," that is, of favoring depth of understanding of scientific concepts over breadth of factual knowledge. Nevertheless, in view of the unanimity seemingly achieved by the mathematicians and mathematics educators and the equally obvious lack of consensus in the sciences, NSTA asked the NRC in early 1991 to coordinate the development of national science education standards.

NRC. The NRC obtained funding for this purpose from NSF and the U.S. Department of Education. It established three working groups: one in content, one in teaching, and one in assessment. Each group comprised 17 to 18 members, including teachers, science education researchers and curriculum developers, and academic scientists and teacher educators. Three additional groups were established: an oversight committee of nearly 40 individuals representing a variety of interests in science education and of organizational affiliations; a Chair's Advisory Committee that included the director of Project 2061, the originator of *SS&C*, and one of the codirectors of the New Standards Project, as well as past and current executive directors of NSTA and NCTM and of other science and science education bodies active in educational reform (e.g., the American Chemical Society, the American Association of Physics Teachers, and the National Association of Biology Teachers); and a small executive editorial committee drawn from the other five groups.

The initial phase of the development of the *National Science Education Standards* (NSES) lasted through the fall of 1993; the standards document *NSES* (NRC, 1996) notes: "During that 18 months [from initiation in May 1992 through fall 1993]. . . . [m]ore than 150 public presentations were made to promote discussion about issues in science education reform and the nature and content of science education standards" (p. 14). An early draft was released in May 1994 and critiqued by a number of focus groups, including several formed by the organizations represented on the Chair's Advisory Committee. A penultimate draft incorporating the resulting comments and suggestions was circulated for "nationwide" review in December 1994; according to the standards document (p. 15) more than 40,000 copies were distributed to some 18,000 individuals and 250 groups.³ The standards document prominently notes the influence of *SFAA* and *Benchmarks* on the individuals involved in the NRC standards development and "gratefully acknowledges its indebtedness to the seminal work by the American Association for the Advancement of Science's Project 2061 and believes that use of the *Benchmarks for Science Literacy* by state framework committees, school and school-district curriculum committees, and developers of instructional and assessment materials complies fully with the spirit of the content standards" (NRC, 1996, p. 15). The limitation of this sanctioning of *Benchmarks* to the *NSES* content standards is significant, as is discussed below.

³ The director of the NRC standards project has written a detailed account of the development of the *National Science Education Standards* (Collins, 1995), including a summary of their contents.

The originator of SS&C was not pleased by the absence of acknowledgment of this NSTA reform initiative. Nevertheless, NSTA has undertaken a series of publications intended to provide compatibility guidelines for NRC's *National Science Education Standards*. The guidelines are intended mainly for teachers to help them make choices about their teaching, materials, sequencing, and assessing student progress. Through providing examples and descriptive dialogue in the guidelines, NSTA hopes to demonstrate how to translate the *Standards*—specifically the content components—into developmentally appropriate activities and investigations. So far, the guidelines for the high school level have been published. Follow-up training institutes for teacher leaders and workshops at NSTA national and regional meetings are planned as well.

The March/April 1996 issue of *Science Education News* (AAAS) carried a front-page article entitled "AAAS, NSTA, and NAS Agree on Reform."⁴ The joint statement by the three organizations outlines the different responsibilities of each. NRC is acknowledged as being responsible for establishing a broad set of standards for science education, building consensus for their acceptance, and tracking progress in their use; NSTA has the role of developing tools, including the guidelines, curricula, and training for the implementation of the standards; and AAAS is credited with setting forth a vision for what it means to be a science literate individual and for providing a coherent set of science learning goals to achieve that vision as well as providing other resources to help teachers and others restructure the science curriculum. This division of labor still leaves some unanswered questions: To what extent do the documents produced by these organizations—and by states, by agencies responsible for student assessment, by textbook publishers, and by others determining the curriculum at least to some extent—agree on the *science content* to be taught? On *instructional strategies*? On the need for *enabling standards*? The next sections of this paper addresses these questions.

The Science Content of *NSES*, the *Benchmarks*, and *The Content Core*⁵

Presumably, in formulating any standards-like document for science education, one needs to identify the concepts considered fundamental in the science disciplines. What should today's students know about physics? About biology? About earth science? About concepts that cut across the disciplines? Which ideas have the most intellectual mileage? Which are most important for personal development, for developing a citizenry able to make informed science-linked decisions, for encouraging further study of science? The amount of content that

⁴ NAS refers to the National Academy of Sciences, an honorific membership organization of scientists. NRC is the working arm of the Academy and provides staff to the volunteer committees convened by the Academy to advise the federal government. The whole complex often is referred to as NAS/NRC.

⁵ Although NSTA's *Content Core* is no longer mentioned prominently in discussions and writings about science education standards (e.g., in state framework documents), this document is of historical importance and continues to undergird ongoing efforts to implement SS&C curricula in the schools. In my view, therefore, it deserves to be included in the content analysis of science standards.

conceivably might be covered in science is potentially enormous and growing more so every year as more new information is generated.

There is philosophical agreement among science educators that the current diet of science courses makes for an overstuffed but undernourishing curriculum and that curricular reform should aim for greater depth of understanding even if at the expense of breadth of knowledge of science facts. Certainly, an attempt has been made in the NRC's *NSES* and the AAAS documents to honor this aim; for example, such staples of the science curriculum as simple machines, series and parallel circuits, phyla of plants and animals,⁶ Ohm's law, and balancing chemical equations have been omitted from both *NSES* and *Benchmarks*. Nevertheless, there are 855 benchmarks in the AAAS document (though a number of these represent building blocks toward a smaller number of complex concepts), 252 major science topics in the sequences presented in NSTA's *Content Core*, and 77 sections representing separate learning goals in *NSES*. If the aim is to focus on what is essential for science literacy and to teach it more effectively, the task given to schools and teachers is hardly less daunting than the present curriculum, overloaded as it is.

A Content Comparison

Comparing the science content the three documents deem essential is not an easy task for two reasons: first, different conceptions of how the science content should be organized and, second, differences on how the topics/standards/benchmarks should be related to grade levels.

Different Organization of Content. With respect to science content, *Benchmarks*, following the original conception in *SFAA*, is the most inclusive as well as the most nontraditional. Only two of its 12 content chapters ("The Physical Setting" and "The Living Environment") and part of a third ("The Human Organism") deal with content generally found in the traditional science curriculum. Content from mathematics, technology, and the social and behavioral sciences is treated in three separate chapters, and the last also in part of the "Human Organism" chapter; three more chapters address the nature of science, of mathematics, and of technology. The remaining three chapters are devoted to "Historical Perspectives," "Common Themes" (a much misunderstood and overused chapter), and "Habits of Mind." The rationale for the inclusivity of *SFAA* and the *Benchmarks* is one of coherence and connectivity, but Project 2061's unorthodox approach has not necessarily been well received by practitioners in fields generally considered separate from the natural sciences. Mathematicians and mathematics educators, for one thing, have developed their own widely acclaimed and accepted standards for K-12 mathematics; for another, they generally see the field of mathematics as much broader than what can be related to or is influenced by science. Moreover, mathematics has enjoyed a privileged status in lower

⁶ Knowledge of phyla is the one topic common to elementary and lower secondary school science in some 50 TIMSS countries, according to an analysis of their national guides and textbooks.

education, traditionally receiving considerable attention in the school curriculum (in fact, rather more so than does science); folding it into the science curriculum is likely to be resisted by mathematics teachers on both substantive and self-interest grounds. Technology, on the other hand, is a newly emerging subject in the U.S. school curriculum, though frequently required in the schools of other industrialized countries. Practitioners and educators in the field are striving to establish technology in K-12 education as separate from science (and from the limited definition of technology common in the science curriculum as being applied science) and indispensable to citizens and workers of the 21st century. Possibly, though, the inclusion of technology in *SFAA* and the *Benchmarks* may give legitimacy to its study in elementary and secondary school. In fact, an effort is under way currently, funded jointly by NSF and NASA, to develop standards for technology education in grades K-12. As to the social and behavioral sciences, their inclusion has generated less controversy, possibly because of their weak presence in the school curriculum, generally limited to civics education, the often disdained "social studies," and an occasional elective in high school such as psychology or economics. On the contrary, social science educators may have welcomed the inclusion of their territory in Project 2061's documents just because social studies is being challenged as a distinctive curriculum subject.

The content standards of *NSES*, contained in chapter 6, are somewhat less bold. They are organized into eight areas, three of which follow the traditional disciplinary divisions of physical, life, and earth/space science, and three of which deal with science and technology, science in personal and social perspectives, and history and nature of science. At times, *NSES* treats two separate chapters in *Benchmarks* as one (e.g., "History and Nature of Science"); sometimes it disaggregates one *Benchmark* chapter into two (e.g., "Physical Science" and "Earth and Space Science"). Both the social sciences and technology are treated as they relate to the natural sciences, not as fields in themselves; mathematics is not treated at all. On the other hand, *NSES* adds as its leading content area "Science as Inquiry," which is partially addressed, though not in the same manner, in the *Benchmark* chapter on "Habits of Mind." An eighth area dealing with "Unifying Concepts and Processes," corresponding to Project 2061's "Common Themes," is briefly described in a separate section; the section covers grades K-12 but is not further detailed by grade level bands, as are the other seven areas comprising the *NSES* content standards.

Most traditional of all is the content organization of NSTA's *Content Core*, which is limited to the four disciplines traditionally studied in grades 6-12: biology, chemistry, earth/space science, and physics. It is worth noting two further characteristics: the weighting given the different disciplines, and the inclusion of many of the traditional topics omitted by both *Benchmarks* and *NSES*. Of the 252 topics listed as core for the seven grades, 31 are in biology, 57 in chemistry, 71 in earth/space science, and 93 in physics. It is not difficult to discern at least one of the reasons, that being the different "grain size" of topic listings in the four disciplines. For example, for biology in grades 11-12, topics include molecular genetics, growth and development (of the living organism), and evolution or theories on the origin of life; for physics at the same grade levels, topics include heat engines and refrigerators, vector addition, parallel circuits, and Doppler effect, though there also are more comprehensive topics listed such as Newton's first and second laws of thermodynamics. The NSTA document states, "*The Content Core* is particularly compatible with the direction, tenets, and themes of the American Association for the Advancement of Science's Project 2061. *The Content Core* quite consciously reflects the 2061 theme that 'less is more'" (p. 9). The actual content of the NSTA document is not consistent

with this statement. Not only does it take an orthodox view of the organization of science content as contrasted to the quite revolutionary approach of Project 2061, but it also includes most of the traditional topics—including all those listed above—that have been omitted in both *NSES* and Project 2061. One possible explanation for this is that the focus is on secondary school, thus warranting greater detail. To be fair, most impartial readers would conclude that *Benchmarks* and *NSES* also did not succeed in making a parsimonious presentation of school science content, despite their omission of some traditional topics.

Table 1 presents a chart relating the science content components of the AAAS, NRC, and NSTA documents at the most global level. A finer-grained analysis presents considerably more difficulty because of the different ways the documents relate content to grade levels and because of the different levels of aggregation of topics within the larger categories of science content.

Organization of Content by Grade Level. Each of the three documents uses different cutpoints for grade levels. *Benchmarks* organizes the thirteen grades K-12 into four levels: K-2, 3-5, 6-8, and 9-12. *NSES* uses three levels: K-4, 5-8, and 9-12; this corresponds to the levels used in the *NCTM Standards* as well as to national and international testing levels.⁷ NSTA organizes the seven grades 6-12 into three bands: 6-8, 9-10, 11-12.

Moreover, *Benchmarks* makes its twelve content areas the superordinate category and discusses each set of grade level benchmarks within a given content topic. *NSES* is organized in the opposite way: the three grade level bands are the superordinate category, with each of the eight standards areas taken up within each grade level band. To help the reader, tables in the rationale section introducing the content standards provide an overview of the eight content areas by grade levels as well as grade levels by content area—the main organization of the *NSES* content standards. The NSTA *Content Core* follows both schematics: It is organized by the four disciplines and topics within them, with a matrix showing what students at the three grade levels are to learn with respect to each subtopic. In the accompanying text, however, descriptions of the subtopics are arranged by grade levels.

The different approaches are illustrated through the following excerpts dealing roughly with the same subject matter: the interdependence of life. Unfortunately, it is not possible in this paper to do justice to the attractive formats and illustrations of *NSES* and the *Benchmarks* and the useful cross-references throughout the chapters in both documents. Nevertheless, the reader may gain a sense of their different styles and differences of both compared to the NSTA document, which was produced with considerably fewer resources.

Excerpts from Benchmarks for Science Literacy

⁷ Both the National Assessment of Educational Progress (NAEP) and the Third International Mathematics and Science Study (TIMSS) focus their student achievement tests on grades 4, 8, and 12; the International Assessment of Educational Progress tested 9- and 13-year-old students in mathematics and science, approximating grades 4 and 8.

The Living Environment: Interdependence of Life (p.115)

It is not difficult for students to grasp the general notion that species depend on one another and on the environment for survival. But their awareness must be supported by knowledge of the kinds of relationships that exist among organisms, the *kinds* of physical conditions that organisms must cope with, the *kinds* of environments created by the interaction of organisms with one another and their physical surroundings, and the complexity of such systems. Students should become acquainted with many different examples of ecosystems, starting with those near at hand.

[Grades K-2 and 3-5 Omitted]

Grades 6 through 8 (p. 117). As students build up a collection of cases based on their own studies of organisms, readings, and film presentations, they should be guided from specific examples of the interdependency of organisms to a more systematic view of the kinds of interactions that take place among organisms. But a necessary part of understanding complex relationships is to know what a fair proportion of the possibilities are. The full-blown concept of ecosystem (and that term) can best be left until students have many of the pieces ready to put in place. Prior knowledge of the relationships between organisms and the environment should be integrated with students' growing knowledge of the earth sciences.

By the end of the 8th grade, students should know that

- In all environments—freshwater, marine, forest, desert, grasslands, mountain, and others—organisms with similar needs may compete with one another for resources, including food, space, water, air, and shelter. In any particular environment, the growth and survival of organisms depend on the physical conditions.
- Two types of organisms may interact with one another in several ways: They may be in a producer/consumer, predator/prey, or parasite/host relationship. Or one organism may scavenge or decompose another. Relationships may be competitive or mutually beneficial. Some species have become so adapted to each other that neither could survive without the other.

Grades 9 through 12 (p. 117). The concept of an ecosystem should bring coherence to the complex array of relationships among organisms and environments that students have encountered. Students' growing understanding of systems in general can suggest and reinforce characteristics of ecosystems—interdependence of parts, feedback, oscillation, inputs, and outputs. Stability and change in ecosystems can be considered in terms of variables such as population size, number and kinds of species, and productivity.

By the end of the 12th grade, students should know that

- Ecosystems can be reasonably stable over hundreds or thousands of years. As any population of organisms grows, it is held in check by one or more environmental factors: depletion of food or nesting sites, increased loss to increased numbers of predators, or parasites. If a disaster such as flood or fire occurs, the damaged ecosystem is likely to recover in stages that eventually result in a system similar to the original one.
- Like many complex systems, ecosystems tend to have cyclic fluctuations around a state of rough equilibrium. In the long run, however, ecosystems always change within climate changes or when one or more new species appear as a result of migration or local evolution.
- Human beings are part of the earth's ecosystems. Human activities can, deliberately or inadvertently, alter the equilibrium in ecosystems.

Excerpts from NSES

Each of the science content standard statements is followed by sections on "Developing Student Understanding" and "Guide to the Content Standard," as illustrated below.

[Grades K-4 are omitted]

Life Science (pp. 155-156)

CONTENT STANDARD C: As a result of their activities in grades 5-8, all students should develop understanding of

- Structure and function in living systems
- Reproduction and heredity
- Regulation and behavior
- Populations and ecosystems
- Diversity and adaptations of organisms

Developing Student Understanding. In the middle-school years, students should progress from studying life science from the point of view of individual organisms to recognizing patterns in ecosystems and developing understanding about the cellular dimensions of living systems. For example, students should broaden their understanding from the way one species lives in its environment to populations and communities of species and the ways they interact with each other and with their environment

Students understand ecosystems and the interactions between organisms and environments well enough by this stage to introduce ideas about nutrition and energy flow, although some students might be confused by charts and flow diagrams. If asked about common ecological concepts, such as community and competition between organisms, teachers are likely to hear responses

based on everyday experiences rather than scientific explanations. Teachers should use the students' understanding as a basis to develop the scientific understanding.

GUIDE TO THE CONTENT STANDARD. Fundamental concepts and principles that underlie this standard include

Populations and Ecosystems (pp. 157-158)

- A population consists of all individuals of a species that occur together at a given place and time. All populations living together and the physical factors with which they interact compose an ecosystem.
- Populations of organisms can be categorized by the function they serve in an ecosystem. Plants and some micro-organisms are producers—they make their own food. All animals, including humans, are consumers, which obtain food by eating other organisms. Decomposers, primarily bacteria and fungi, are consumers that use waste materials and dead organisms for food. Food webs identify the relationships among producers, consumers, and decomposers in an ecosystem.
- For ecosystems, the major source of energy is sunlight. Energy entering ecosystems as sunlight is transferred by producers into chemical energy through photosynthesis. That energy then passes from organism to organism in food webs.
- The number of organisms an ecosystem can support depends on the resources available and abiotic factors, such as quantity of light and water, range of temperatures, and soil composition. Given adequate biotic and abiotic resources and no disease or predators, populations (including humans) increase at rapid rates. Lack of resources and other factors, such as predation and climate, limit the growth of populations in specific niches in the ecosystem.

Life Science (p. 181)

CONTENT STANDARD C: As a result of their activities in grades 9-12, all students should develop understanding of

- The cell
- Molecular basis of heredity
- Biological evolution
- Interdependence of organisms
- Matter, energy, and organization in living systems
- Behavior of organisms

Developing Student Understanding. Students in grades K-8 should have developed a foundational understanding of life sciences. In grades 9-12, students' understanding of biology will expand by incorporating more abstract knowledge, such as the structure and function of DNA, and more comprehensive theories, such as evolution. Students' understandings should

encompass scales that are both smaller, for example, molecules, and larger, for example, the biosphere.

GUIDE TO THE CONTENT STANDARD: Fundamental concepts and principles that underlie this standard include

The Interdependence of Organisms (p. 186)

- The atoms and molecules on the earth cycle among the living and nonliving components of the biosphere.
- Energy flows through ecosystems in one direction, from photosynthetic organisms to herbivores to carnivores and decomposers.

- Organisms both cooperate and compete in ecosystems. The interrelationships and interdependencies of these organisms may generate ecosystems that are stable for hundreds or thousands of years.
- Living organisms have the capacity to produce populations of infinite size, but environments and resources are finite. This fundamental tension has profound effects on the interactions between organisms.
- Human beings live within the world's ecosystems. Increasingly, humans modify ecosystems as a result of population growth, technology, and consumption. Human destruction of habitats through direct harvesting, pollution, atmospheric changes, and other factors is threatening current global stability, and if not addressed, ecosystems will be irreversibly affected.

Excerpts from The Content Core

The NSTA document is organized by the type of matrix shown in Table 2 for a part of the biology sequence (reproduced from p. 35 of *The Content Core*), followed by explanatory text for each grade level band, excerpts of which follow.

Biology 6-8 (pp. 37-38)

THE BIOLOGICAL PLANET

The third biology content organizer is the biological planet.

Interactions

At the 6-8 grade level, activities exploring the interactions of organisms and the environment should be confined to the physical environment: limiting factors, temperature, moisture, and light. These activities also should differ from earlier activities that explored how individual organisms respond to specific stimuli. Students will consider interrelationships between organisms, including humans and other organisms, in later grades.

Biology 9-10 (pp. 43-44)

THE BIOLOGICAL PLANET

Having studied the biological planet descriptively in grades 6-8, students in grades 9-10 examine the dynamics of organisms with their environments and with each other. Further, students investigate patterns of energy flow and build on their knowledge of how matter cycles in the environment.

Interactions

Interrelationships between organisms: Students should observe the interrelationships between individual organisms such as predation, symbiosis, mutualism, and parasitism. It is important that students look at ecosystems, in which all these interrelationships operate simultaneously. The role of rare and endangered species as indicators of the transformation of, or health of, ecosystems can be addressed. Oceans and species-rich tropical rainforests are appropriate examples of the Earth's remarkable biodiversity. Cross disciplinary explorations of the role of fire, climate changes, or volcanism on ecosystems over time also can be conducted here.

Activities should include exploration of a local "environment," or one established within the school. Students should observe interactions between populations and consider phenomena such as plant succession and pioneering populations.

Effects of humans on the environment: The 9-10 grade curriculum should address how individual organisms alter the environment. These effects are not caused exclusively by human populations. Dutch Elm disease or killer bees, or chemical pollutants in water resources can initiate a discussion on the complex effects of humans on the environment, a recurring topic throughout secondary school science. The biological implications of human interaction can focus activities here: the effect of human populations on the extinction of other animal and plant species and the accelerated change to the environment leading to habitat loss or the creation of new habitats. Case studies, with the engagement of related ethical and social issues, are productive approaches to this topic.

Biology 11-12 (pp. 47-48)

THE BIOLOGICAL PLANET

Students should consider evolution as the great unifying principle of biology. They should observe that while other non-living systems are said "to evolve," these systems do not operate like biological evolution.

Up to this point, students have encountered descriptively the components of the biological planet and the interrelationships that address issues of adaptation. Evolution, the theories of its mechanism and its various forms and patterns, should focus study in grades 11-12.

Interactions

Evolution: Students should explore and compare theories on the origin of life, dealing with both chemical and biological evolution. In other words, they should consider both the formation of chemical compounds identified with living systems and theories that attempt to explain how species change over time. Controversy over whether species are fixed or mutable has gone on since well before Darwin, and scientists have proposed numerous mechanisms for explaining species change. In order to avoid the stultifying effects of lecturing about these ideas, or only reading about them, students should discuss and test theories about the mechanism of evolution. For instance, the question of why Lamarck's explanation is inadequate to account for certain observations should require students' analysis, rather than the teacher's automatic dismissal of it. Natural selection: Central to understanding evolution is understanding natural selection as its primary mechanism. Activities or discussion should consider the elements of natural selection, the accumulated evidence compiled by Darwin, and the theory's modification as scientists better understood inheritance. Other current theories of the mechanisms for evolution, such as punctuated equilibria, endobiocytosis, and neoDarwinism, should be studied in broad outline and the strengths and weaknesses of each analyzed.

Changing populations: Curricula should give some attention to population genetics and the population as the evolving unit.

Concordance Between Benchmarks and NSES

Obviously, *Benchmarks* is by far the most inclusive in its treatment of science content, *NSES* moderately so, and *The Content Core* the most traditional, so that there are a number of areas missing in *NSES* that can be found in the *Benchmarks*, and a number in both *Benchmarks* and *NSES* that are not included in *The Content Core*, as illustrated at a macro level in Table 1. Beyond that, it still is of interest to compare the subtopics included in the overlapping chapters of *Benchmarks* and standards of *NSES*.⁸ It is beyond the scope of this paper to do a detailed content analysis comparing the treatment of the major science topic areas that the two documents have in common.⁹ Nevertheless, a careful reading of *Benchmarks* and *NSES* leads to the conclusion that, for most areas of school science, there is a great deal of overlap in topic coverage. This is

⁸ Because it deals only with secondary school, takes a more familiar approach to what students should learn in science at that level, and is no longer generally considered a key "standards" document, I have omitted *The Content Core* from the more finely grained analysis.

⁹ Such an analysis has been done by AAAS comparing *NSES* to the *Benchmarks*; the analysis is not publicly available, however.

particularly true of the traditional science subjects taught in school: the physical, life, and earth/space sciences. *NSES* appears to go into somewhat greater detail in chemistry; disease issues receive more attention in *NSES*, whereas the *Benchmarks* include material on learning, reasoning, and mental health. There also is some difference in the treatment of the three themes that both Project 2061 and *NSES* see as cutting across the sciences; for example, somewhat different aspects of systems are emphasized, and constancy and change are treated somewhat differently.

Aside from material not treated at all in *NSES* (notably mathematics, except for measurement), the greatest differences occur, not unexpectedly, in areas where either the *Benchmarks* or *NSES* devotes more space to a set of topics, as is the case for the *NSES*'s single History and Nature of Science standard, given two separate chapters in *Benchmarks*, or the single *NSES* standard on Science and Technology, also given two chapters in *Benchmarks*. On the other hand, *NSES* has a whole standard on Inquiry, treated in *Benchmarks* as part of the Nature of Science chapter. The difference here is more than one of coverage. The Inquiry standard of *NSES* (pp. 121, 143, 173) is stated in the same language for *NSES*'s three grade bands:

"As a result of activities in grades K-4 (5-8, 9-12), all students should develop

- Abilities necessary to do scientific inquiry
- Understanding about scientific inquiry"

The specifications of this standard are, of course, different for the three grade bands; for example, at the primary level, students acquiring abilities necessary to do scientific inquiry should be able to "ask a question about objects, organisms, and events in the environment" (p. 122); at the middle school level, they should be able to "identify questions that can be answered through scientific investigation" (p. 145); at the secondary level, they should be able to "identify questions and concepts that guide scientific investigation" (p. 175). These statements and the several others that specify this standard make it quite clear that the intent is to have students, as a result of their K-12 science education, *be able to carry out* scientific investigations. Project 2061, striving for science literacy, holds it important that students *understand* the process of scientific inquiry to become informed consumers of science information; *Benchmarks* adds to this the need to understand the scientific enterprise.

Supporting Standards in *NSES*¹⁰

Both *Benchmarks* and *The Content Core* have included chapters discussing the rationale for their recommendations and considerations regarding their implementation. *NSES* goes considerably further. In addition to the teaching and evaluation vignettes that illustrate a number of the standards at the three grade level bands, NRC has developed standards for Science Teaching (six standards), Professional Development of Teachers of Science (four standards), Assessment (five standards), Science Education Program Standards (six standards) and Science Education System Standards (seven standards). As are the content standards, each of these additional five standards areas is illustrated by at least one vignette; otherwise, the standards are specified in fairly parsimonious form, as the following examples from three of the standards areas show.

Excerpts from Science Teaching Standards

Teaching Standard A (p. 30): Teachers of science plan an inquiry-based science program for their students. In doing this, teachers

- Develop a framework of yearlong and short-term goals for students.
- Select science content and adapt and design curricula to meet the interests, knowledge, understanding, abilities, and experiences of students.
- Select teaching and assessment strategies that support the development of student understanding and nurture a community of science learners.
- Work together as colleagues within and across disciplines and grade levels.

Develop a Framework of Yearlong and Short-Term Goals for Students. All teachers know that planning is a critical component of effective teaching. One important aspect of planning is setting goals. In the vision of science education described in the *Standards*, teachers of science take responsibility for setting yearlong and short-term goals; in doing so, they adapt school and district program goals, as well as state and national goals, to the experiences and interests of their students individually and as a group.

Once teachers have devised a framework of goals, plans remain flexible. Decisions are revisited and revisited in the light of experience. Teaching for understanding requires responsiveness to students, so activities and strategies are continuously adapted and refined to address topics arising from student inquiries and experiences, as well as school, community, and national events. Teachers also change their plans based on the assessment and analysis of student achievement and the prior knowledge and beliefs students have demonstrated. Thus, an inquiry

¹⁰ I have labeled the five other sets of standards in *NSES* "supporting" standards, though NRC likely considers them of equal importance to the content standards. My reasons for doing so are (a) that they are in support of the ultimate goal of student learning in science and (b) that while the chapters for each of these standards vary from 20 to 30 pages, the chapter defining the content standards is over 100 pages long.

might be extended because it sparks the interest of students, an activity might be added because a particular concept has not been understood, or more group work might be incorporated into the plan to encourage communication. A challenge to teachers of science is to balance and integrate immediate needs with the intentions of the yearlong framework of goals.

During planning, goals are translated into a curriculum of specific topics, units, and sequenced activities and help students make sense of their world and understand the fundamental ideas of science. The content standards, as well as state, district, and school frameworks, provide guides for teachers as they select specific science topics. Some frameworks allow teachers choices in determining topics, sequences, activities, and materials. Others mandate goals, objectives, content, and materials. In either case, teachers examine the extent to which a curriculum includes inquiry and direct experimentation as methods for developing understanding. In planning and choosing curricula, teachers strive to balance breadth of topics with depth of understanding.

Excerpts from Assessment in Science Education Standards

Assessment Standard A (pp. 78-79): Assessments must be consistent with the decisions they are designed to inform.

- Assessments are deliberately designed.
- Assessments have explicitly stated purposes.
- The relationship between the decisions and the data is clear.
- Assessment procedures are internally consistent.

The essential characteristic of well-designed assessments is that the processes used to collect and interpret data are consistent with the purpose of the assessment. That match of purpose and process is achieved through thoughtful planning that is available for public review.

Assessments Are Deliberately Designed. Educational data profoundly influence the lives of students, as well as the people and institutions responsible for science education. People who must use the results of assessments to make decisions and take actions, as well as those who are affected by the decisions and actions, deserve assurance that assessments are carefully conceptualized. Evidence of careful conceptualization is found in written plans for assessments that contain

- Statements about the purposes that the assessment will serve.
- Descriptions of the substance and technical quality of the data to be collected.
- Specifications of the number of students or schools from which data will be obtained.
- Descriptions of the data-collection method.
- Descriptions of the method of data interpretation.
- Descriptions of the decisions to be made, including who will make the decisions and by what procedures.

Excerpts from Science Education System Standards

System Standard C (pp. 231-232): Policies need to be sustained over sufficient time to provide the continuity necessary to bring about the changes required by the *Standards*.

Achieving the vision contained in the *Standards* will take more than a few years to accomplish. Standard C has particular implications for organizations whose policies are set by elected or politically appointed leaders. New administrations often make radical changes in policy and initiatives and this practice is detrimental to education change, which takes longer than the typical 2- or 4-year term of elected office. Changes that will bring contemporary science education practices to the level of quality specified in the *Standards* will require a sustained effort.

Policies calling for changes in practice need to provide sufficient time for achieving the change, for the changes in practice to affect student learning, and for changes in student learning to affect the scientific literacy of the general public. Further, policies should include plans and resources for assessing their affects over time. If school-based educators are to work enthusiastically toward achieving the *Standards*, they need reassurance that organizations and individuals in the larger system are committed for the long term.

System Standard E (pp. 232-233): Science education policies must be equitable.

Equity principles repeated in the introduction and in the program, teaching, professional development, assessment, and content standards follow from the well-documented barriers to learning science for students who are economically deprived, female, have disabilities, or from populations underrepresented in the sciences. These equity principles must be incorporated into science education policies if the vision of the standards is to be achieved. Policies must reflect the principle that all students are challenged and have the opportunity to achieve the high expectations of the content standards. The challenge to the larger system is to support these policies with necessary resources.

The addition of these standards supporting the content standards makes the *NSES* document serve the functions of the three separate NCTM standards documents on curriculum and evaluation, teaching, and assessment. Project 2061 is working on its own implementation documents, including *Designs for Science Literacy* (in press), which will provide guidance on how to reshape the K-12 curriculum around the goals of *SFAA* and the *Benchmarks*; and *Blueprints for Reform*, a series of twelve documents laying out how various parts of the education system need to change to support a curriculum for science literacy (AAAS, 1995). NSTA for its part, as noted, plans to publish guidelines to help teachers and local administrators implement science curricula compatible with the *NSES* recommendations. Both organizations might argue, with justification, that their implementation activities go beyond the publication of documents to action programs involving teachers, schools, and other critical actors and organizations directly. NRC, too, is planning a number of information and dissemination

activities intended to further implementation of *NSES*, perhaps analogous to its earlier establishment of the Mathematical Sciences Education Board (MSEB) in support of NCTM's mathematics standards. NRC is unlikely, however, to engage in project activities taking it directly into schools and into training of teachers and administrators, as is the case for Project 2061 and NSTA. Even the advocacy activities of the MSEB and those now planned for *NSES* are somewhat unusual for the Academy complex, of which NRC is a part.

Assessment Standards

The ultimate short-term measure of what students have learned in science is through assessment of their knowledge and performance. Teachers, administrators, parents, and the students themselves look toward science tests as providing the critical guidance on what is important to teach and learn. In recognition of the importance of assessment, both NCTM and *NSES* have developed assessment standards. As with curriculum itself, however, standards are not enough; they must be translated into day-by-day lessons in the case of curriculum and into test items or performance exercises in the case of assessment. Three major assessment efforts are of note as related to standards-based reform in science education.

The New Standards Project

This project was created to develop assessment systems that states and local districts could use to "measure their students' progress toward meeting national standards that are internationally benchmarked" (National Center on Education and the Economy [NCEE], 1995a, p. 2). Three components are planned: performance standards, on-demand examinations, and a portfolio system. In science, only the first of these, the performance standards, have been developed. The performance standards are intended to (a) establish high standards for all students, (b) be rigorous and world class, and (c) be useful, developing what is needed for citizenship, employment and life-long learning (p. 3). Other criteria the project cites for its performance standards include both importance and parsimony in judging what to emphasize, manageability with respect to time, flexibility to permit local control and adaptation, and being based on broad consensus building.

The performance standards for science are published in three volumes (NCEE, 1995a, b, c) covering elementary school, middle school, and high school. Each volume contains performance standards for English language arts, mathematics, and applied learning as well as for science. The project has set out its own content standards; for science, these include physical science concepts, life science concepts, earth and space science concepts, scientific connections and applications, scientific thinking, scientific tools and technologies, scientific communication, and scientific investigation. Each of these standards is defined by means of three to six bullets; there is some variation in this language from volume to volume to reflect increasing expectations regarding students' science knowledge and sophistication of performance. Brief examples of performance that may demonstrate understanding also are given. The following excerpts from the *Performance Standards* deal with the Earth and Space Sciences Concepts as they are given in the three volumes.

ELEMENTARY SCHOOL
(from NCEE, 1995a, p. 62)

Earth and Space Sciences Concepts

The student understands:

- properties and uses of Earth materials, including rocks, soils, water, and gases;
- patterns, cycles, seasons, time, weather, and Earth motion;
- change over time, for example, erosion.

Examples of performances that may demonstrate understanding include:

- Δ identifying features of the school building that are related to the weather; explaining what would change inside the classroom if they were not present;
 - Δ keeping a record of the shape of the moon for several months; predicting what will happen in the next week;
 - Δ collecting information from a weather station and using the information to describe the changes from fall to winter (see also Mathematics Standards 1 and 4; Applied Learning Standard 1);
 - Δ writing a story that tells what happens to a drop of water when it goes from a lake to a river.
-

MIDDLE SCHOOL

(from NCEE, 1995b, p. 48)

Earth and Space Sciences Concepts

The student understands:

- Earth's systems, including crustal plates and land forms; rock cycle, water cycle; weather and oceans;
- Earth's history, especially change over time, erosion, movement of plates, fossil evidence;
- Earth in the Solar System, including day, year; sun, planet; gravity, energy;
- natural resource management.

Examples of performances that may demonstrate understanding include:

- Δ explaining why earthquakes, volcanoes, and sea-floor spreading have a common cause;
 - Δ writing a story about the experiences of a water molecule as it travels the globe;
 - Δ predicting what happens to the reading on a bathroom scale while riding in an elevator and explaining your observations;
 - Δ using the concept of gravity to explain why people can jump higher on the moon than they can on Earth;
 - Δ developing an algorithm to tell whether the Moon is waxing or waning;
 - Δ completing the Geology Project (Girl Scouts of America) or earning the Astronomy Merit Badge (Boy Scouts of America) and explaining how it helped you to understand and Earth sciences concept.
-

HIGH SCHOOL

(from NCEE, 1995c, p. 56)

Earth and Space Sciences Concepts

The student understands:

- Earth's systems, including the Sun, radioactive decay, gravitational energy; weather and climate;
- origin and evolution of the Earth system, in particular, estimating geologic time, age of life forms;
- forces that shape the Earth; that is, processes and observable results;
- natural resource management.

Examples of performances that may demonstrate understanding include:

- △ providing an orientation to the climate of the local region to a newcomer; explaining today's weather in that context;
- △ explaining why people can jump higher on the Moon than they can on Earth;
- △ explaining the relationship between gravity and energy;
- △ analyzing the risk of natural disasters in the local region and making recommendations for actions that can be taken to mitigate the damage;
- △ conducting a study of the geology of an area near the school; describing the likely history of the region, using observations and reference materials.

The contribution of the New Standards Project is not in the formulation or detailing of standards per se, but in providing elaborated student work samples and commentary on selected performance tasks. In science, there are nine of these at the elementary level, nine at the middle school level, and eight at the high school level. The tasks either have been specially designed or drawn from existing sources, such as classroom projects, 4-H projects (for elementary school), and entries to the California Golden State Examination Science Portfolio (for high school). Some tasks address such common science topics as the ecosystem of an aquarium in elementary school; buoyancy, light reflection, classification of seeds, and earth in the solar system in middle school; and erosion and pollution in high school. There is no attempt to cover all the bullets specifying any of the standards. At least half the tasks included at each level are quite familiar including, for example, at the middle school level, the paper towels task that has been studied in some detail by Shavelson et al. (1991), the classification of seeds task, the spot remover task, and the phases of the moon task.

Each task is given one or two pages in an 11" x 17" volume, occasionally three or four pages in the high school volume. The large format allows space for a short description of the task, which of the standards the task addresses (i.e., the content knowledge and performance being expected), reproductions of actual student work, commentary on that work and some judgment of it, comments on the suitability of the task when warranted, and "international benchmarks." The last generally consist of a one-sentence (occasionally one-paragraph) quote from another country's curriculum guidelines. For example, the international benchmark for the photosynthesis laboratory, a high school task, quotes the Australian science guidelines as expecting students to explain "how living things obtain, store and transport nutrients, transform energy, and manage wastes. [This is] evident, for example, when students investigate the way green plants use sunlight to produce simple sugars in photosynthesis" (NCEE, p. 61, 1995c). As a benchmark, this is very broad and hardly provides much purchase on how one might judge whether American students are performing up to world class standards.

Likely of most interest to teachers, given the challenge to bring all students to high levels of performance, are the commentaries on how a particular student performance is being evaluated by the authors of the New Standards volume: what level of quality the work indicates and how the student might have gone further. For example, for a high school task on the density of sand, the commentary states in part (NCEE, 1995c, p. 58):

Science evident in this student response

In this entry, a pair of students were asked to devise and carry out a method for determining the density of sand with air around the sand granules and the density of sand alone. Because they were investigating density, they were required to demonstrate understanding of parts of the following Science standards:

Standard 1, Physical Sciences

Concepts—structure and properties of matter;

Standard 5, Scientific Thinking;

Standard 6, Scientific Tools and Technologies.

Physical Sciences Concepts

The work shows clear evidence for understanding the concept of density, e.g., “The equation for density is mass divided by volume.” The inverse relationship that exists between density and volume is given in the statement, “Since the volume of the sand [with] air was larger, it had a lower density.” This flexibility with a ratio concept of density would not be expected at the middle school level, where an understanding of the concept in concrete, physical terms is expected. The high school level understanding is further evident in the statement, “Since density is an intense property, the difference in sample sizes among the other groups should not have affected the results.” And the discussion of the real world applications of this understanding (see item #3 on the Self-reflection Sheet) shows that the student can generalize the situation from the immediate context.

Scientific Thinking

Scientific Tools and Technology

The assignment required that the students develop an appropriate procedure. . . .

There is a clear attention to accuracy and precision throughout the work, e.g., “we devised a more accurate plan of weighing the sand within the cup. . . .” The use of the graduated cylinder for the dry sand and then the use of the same equipment for the water displacement method shows attention to accuracy as well.

The comparison of one group to four others was further evidence of the check for accuracy tied to Scientific Tools and Technologies. Three of the five groups had results which were similar. . . . The suggestion of all groups double checking their measurements and calculations is consistent with the quality of work required by Scientific Tools and Technologies.

Going Beyond

This work shows evidence for part of the standard for Physical Sciences Concepts but would need to be accompanied by work of similar quality with chemical reactions, forces and motion, and energy to meet the standard. Similarly, additional work would be needed to meet the standards for Scientific Thinking and Scientific Tools and Technologies, in particular, alternative explanations and multiple data sources.

Teachers and science educators involved in evaluating student performance on the sorts of tasks the New Standards Project has collected will deepen their understanding of how to assess what students are learning in their science classes. This ought to help teachers in developing assessments aligned to the kinds of goals espoused by *NSES* and the other standards-setting efforts; possibly, it will also provide them with insights on how to improve science instruction to attain the goal of high achievement for all students. Engaging teachers around assessment and the evaluation of student performance has proved a highly effective strategy for staff development, both here and in England, as teachers start to reflect on their expectations for student performance and on criteria for judging different levels of quality.

The National Assessment of Educational Progress (NAEP)

Because NAEP prepares reports called *The Nation's Report Card*, because it has become a high-stakes test for states as their students' scores are compared to each other and to national norms, and because it now reports out not only students' aggregate achievement levels but also their performance according to predefined standards, this national and state-by-state assessment is looked to as the closest means available for assessing the extent to which students learn the content of various subject-matter standards. Some historical background is useful to understand how NAEP came to play this role.

NAEP was begun in 1969 to monitor student achievement in core subjects at three age/grade levels (ages 9, 13, 17/grades 4, 8, 12). The assessment takes place every two years in selected subjects; it uses nationally representative samples of some 25,000 students (5,000-8,000 at each grade level) drawn independently for each subject. Over time, three crucial developments have made NAEP a somewhat higher-stakes assessment than was originally intended.

First, state assessments, as well as the national assessments, have been conducted, starting with mathematics in grade 8 in 1990 and grades 4 and 8 in 1992; in 1994, reading in grade 4 was added to the two mathematics state assessments. Although state participation is voluntary, most states have been participating, making state-by-state comparisons possible. There have been demands to extend NAEP to make district comparisons possible, and even to extend NAEP to all students instead of representative national and state samples, but this is an unlikely development at present, if for no other reason than fiscal constraints. Hence, NAEP remains a low-stakes assessment for individual students and teachers.

Second, the National Assessment Governing Board (NAGB) added the requirement that NAEP set desired performance levels for purposes of reporting results, rather than just reporting scale scores (or percent correct). Currently, NAEP sets three performance levels for each of the three grades tested: Basic, Proficient, and Advanced; there are four reporting categories since "Below Basic" also is reported. The determination of these performance levels has been difficult and controversial; it obviously ties into judgments of what standards should be applied to student performance.

This brings up the third development: the concern that the NAEP assessment be aligned with reform efforts and nationally developed standards in any subject area being assessed. For example, earlier science assessments were criticized for being almost exclusively devoted to testing the memorization of facts without concerns for coherent understanding. The push toward aligning NAEP with current visions of, say, mathematics or science education as exemplified in the NCTM *Standards*, *NSES*, and *Benchmarks* makes for difficult compromises for the developers of the NAEP content frameworks that guide the development of the tests and for developing the tests themselves. On the one hand, NAEP is asked to be forward-looking, to provide tests that embody or at least are consonant with standards that are often quite visionary, and thus to become a driver of reform; on the other hand, a certain realism about what actually is being taught in schools has to inform both the guiding NAEP frameworks and the assessments themselves, else a large majority of students is likely to end up in the "Below Basic" category. Also, trend lines (comparing student achievement over the years) cannot be maintained if the assessment shifts too radically, although technical fixes are available to ameliorate this problem.

There has been no science assessment since 1990. A framework was developed for a planned 1994 science assessment that was postponed until 1996. Originally, state science assessments were proposed for grades 4 and 8 in 1996, but funds were available only for a grade 8 state-by-state assessment as well as for the national assessment. The framework guiding the construction of the assessment, subsequently revised (NAGB, 1996), is forward looking in that it recommends that at least 30 percent of the assessment, as measured in student response time, should be devoted to hands-on performance exercises, that 50 percent should be devoted to open-ended items, and that multiple-choice items should comprise no more than 50 percent of the assessment. (Some of the responses to the performance exercises could be in multiple-choice format.) The NAEP science framework document makes other recommendations to bring the 1996 NAEP assessment in line with current reform directions in science education, but there have been limits on the extent to which the recommendations were incorporated in the tests, which were administered in spring 1996.

Some key individuals involved in the development of *NSES* and the *Benchmarks* also served on the committees responsible for the current NAEP science framework. Nevertheless, its organization of science content is quite traditionally structured according to the matrix reproduced in Figure 1. The three fields of science named on the matrix are divided into three to four categories each, each of which is further subdivided into three to seven subtopics, making 45 subtopics in all.

Two aspects of the framework are noteworthy: The first is the inclusion of "Nature of Science" (which is intended to include both science and technology) and the three themes cutting across all three fields of science. These additions to the traditional fields of science taught in school are

quite consonant with *NSES* and the *Benchmarks*. The second noteworthy aspect is the set of "Knowing and Doing" categories, particularly two of the subcategories of knowledge included under "Conceptual Understanding" ("propositions about the nature, history and philosophy of science" and "kinds of interactions between and among science, technology, and society") and the "Practical Reasoning" category intended to address the ability of eighth- and twelfth-graders to apply their science knowledge to everyday problems and, as they mature, to societal problems linked to science and technology. It remains to be seen to what extent the NAEP science tests administered in 1996 will address the more innovative aspects of the NAEP framework, which as yet have a less secure psychometric base than do multiple-choice items and also tend to be more costly to administer and score.

The Third International Mathematics and Science Study (TIMSS)

This is the largest study of student achievement in science (and mathematics) ever undertaken. Not only did more than 45 countries participate, but a wealth of information on contextual variables has been collected to array against the testing outcomes. The achievement tests were administered in 1994-95 to students in five grades, encompassing approximately the same three levels as NAEP. The students, as well as their teachers and the principals of their schools, also were asked to respond to questionnaires about their backgrounds, attitudes, experiences, and teaching practices. In addition, a major study of the intended curriculum in each country was undertaken, based on detailed analyses of national curriculum guidelines (or regional guidelines in the case of decentralized systems) and textbooks in most common use. Curriculum frameworks for mathematics and science were created to provide a common language system for all the components of TIMSS, including the construction of the tests and the curriculum analysis work. Each framework has three aspects: content, performance expectations, and perspectives.

The subcategories of the **Perspectives Aspect** are the same for both science and mathematics and deal with attitudes toward science, careers and participation in science, increased interest in scientific topics, and habits of mind. For science, the **Content Aspect** and the **Performance Expectations Aspect** are configured as follows (Robitaille et al., 1993):

· **Content Aspect**

- Earth sciences
- Life sciences
- Physical sciences
- Science, technology, and mathematics
- History of science and technology
- Environmental and resources issues
- Nature of science
- Science and other disciplines

· Performance Expectations Aspect

Understanding

Theorizing, analyzing, solving problems

Using tools, routine procedures

Investigating the natural world

Communicating

The eight science content categories are further subdivided to provide 76 subcategories. Given that more than 45 countries had to agree to the science framework, the subcategories for the earth, life, and physical sciences are reasonably traditional; the other five content categories comprise—broadly stated—much of parallel material in *NSES* and the *Benchmarks*. The performance expectations categories are subdivided into 20 subcategories, some of which also reflect current reform directions in this country and abroad. For example, the "understanding" category includes understanding "simple information," "complex information," and "thematic information"; the "investigating the natural world" category includes "identifying questions," "designing investigations," "conducting investigations," "interpreting data," and "formulating conclusions."

One of the innovative aspects of the TIMSS frameworks is the possibility of double-coding, that is, a given test item or lesson from a textbook might deal with one or more subcategories in the life sciences and with environmental issues as well, or with the physical sciences and the history of science. By identifying and coding all the appropriate subcategories, one is able to identify cross-connections in intended science instruction or in the test items. Moreover, throughout the TIMSS tests and the analyses of country guidelines and textbooks, every content subcategory and associated code that has been identified for a given test item or guideline/textbook passage has been linked to one or more performance expectation subcategory and code. This makes possible relational analyses between specific subject matter and expectations for what students should know and be able to do regarding given science concepts.

The results of TIMSS, particularly those stemming from the curriculum analyses and the student test data, should provide the most detailed and complete information yet available on what other countries think is important to teach in science, how they think it should be taught, what they expect of students by way of science knowledge and the ability to do science, and what students actually have learned—in other words, the information needed to construct "world-class standards." The first TIMSS results will deal with the intended curriculum in science and mathematics for students in primary and secondary school in the 45 countries, drawn from analyses of the countries' national and regional curriculum guidelines and from textbooks in common use in each country; the relevant reports became available in October 1996. The test results for 13-year-olds (eighth-graders in U.S. terms) became available in late November 1996, accompanied by selected contextual information.

Analysis of University Entrance Examinations

TIMSS did not have the resources to include in its curriculum work the analyses of examinations used in various countries, another important determinant of what gets taught and learned. Hence, the National Center for Improving Science Education undertook such an analysis, using the TIMSS frameworks. The examinations analyzed were university entrance examinations from

seven countries, including England and Wales, France, Germany, Israel, Japan, Sweden, and the U.S. University examinations were selected since they are given in most countries at approximately the same level of education (the completion of secondary schooling) and represent the most rigorous expectations of student performance at that level—one aspect to be considered in the development of standards that are to be informed by what other countries expect of their students. (Advanced Placement [AP] Examinations were analyzed for the U.S., rather than the Scholastic Assessment Test [SAT] or the College Board Achievement or American College Testing examinations, none of which—though taken by many students seeking college admission—are sufficiently rigorous to compare to the university entrance examinations of other countries.)

Examinations in biology, chemistry, and physics were analyzed as well as those in mathematics. The results are reported in Britton and Raizen (1996). The authors conclude that it is not appropriate to rank the examinations from the seven countries according to difficulty, because such ranking would vary according to the criteria used, e.g., length of the exam, whether students can choose among questions, type of item predominating in the exam, breadth and depth of the exam, and what student performances are expected. Also, countries' educational systems are organized differently, as are their examination systems, including incentives for taking the exams, grading and scoring, pass rates, and extent of coaching for the exams.¹¹ Nevertheless, some interesting findings emerged from the analyses:

- The AP exams use far more multiple-choice items (about 50%) than do the exams of any other country. French and German exams use no multiple-choice items at all.
- Laboratory practicals are part of the exams in England/Wales and Israel, but not in the other countries.
- All chemistry exams except for the U.S. AP exams devote considerable attention to organic chemistry and some attention to industrial applications of chemistry; these are major differences in what students abroad are expected to learn in this field compared to U.S. students.
- Physics exams rarely cast problems in real-world contexts—despite reform recommendations to the contrary in most countries. This is also true of the mathematics exams, which treat the subject as an abstract discipline unconnected to any physical phenomena.
- The countries with the greatest emphasis on assessing scientific experimentation are England/Wales, France, and Israel. Only Japan's exams require mathematical reasoning to any extent; the other countries' mathematics exams focus on routine procedures and problem-solving.

¹¹ All the exams (except the Swedish and Israeli ones) are being published jointly by the Center and the AFT; the biology exams published in 1994 and the chemistry/physics exams (one volume) published in 1996 are available through AFT; the mathematics exams will be published in 1997. The publications will allow teachers and other educators to judge the difficulty of the exams from other countries for themselves.

Some of the more interesting findings have to do with the use of the exams rather than differences and similarities among the exams themselves. Even given the different incentives, it is noteworthy that only about 6.6% of U.S. students take AP exams, while roughly 25%-50% percent of all students in the other nations take and pass these types of advanced exams. This generally comes as a shock to Americans who still think of the European and Japanese systems as elite and exclusionary, but this has changed considerably in the last two or three decades. Indeed, "the great majority of college-bound students in countries other than the United States *must* take and pass some advanced subject-specific examinations. . . . In France, Germany, and Israel, academically oriented students who do not seek further education still take these examinations because passing them is a prestigious credential in their societies" (Britton & Raizen, 1996, p. 201). Moreover, in most of the countries except the U.S., students must take these examinations in several fields, varying from three subjects in England/Wales to seven or eight subjects in France. Only Japan and the U.S. charge students for taking the exams, which may deter some students in the U.S., but apparently not in Japan, where over 50% of the students take them.

The conclusion seems inescapable that, even without ranking the examinations according to difficulty, other countries are bringing a greater percentage of their students to a high level of achievement than is the U.S.

Science Standards in the States

The state level counterpart closest to national education standards are curriculum documents designed to guide school instruction across a given state. Historically, states have varied greatly regarding the weight given to such documents, from New York's (Board of) Regents syllabi followed by every high school in the state to states—including Maine, Alaska, and Hawaii—that currently are developing curriculum frameworks for the first time. The construction of state curriculum objectives is not a new phenomenon, however; it was spurred on in the 1970s by the advent of competency testing and the increased use of norm-referenced standardized tests and associated examinations. According to a survey conducted by the Council of Chief State School Officers (CCSSO) in 1987 (Blank & Espenshade, 1988), even at that time, 38 states had a curriculum guide or state objectives in science.

Several recent developments have served to accelerate the movement toward state curriculum frameworks¹² for those states that did not previously have them or toward revision of existing ones. First, the publication of the NCTM *Standards* led to states' revision of their own mathematics frameworks. (Many state mathematics curriculum specialists had been heavily involved in drafting and reviewing the national standards.) The advent of the AAAS *Benchmarks*, NSTA's *SS&C Content Core* monograph, and several drafts of the NRC *Standards* similarly renewed interest in state science frameworks.

Further incentives were provided by the federal government: NSF made it clear that strong state standards linked to the national standards were one of the drivers of reform for the states that had

¹² Though states have varying objectives and titles for state documents intended to serve as curriculum guides, we follow general current usage and, in the rest of this section, refer to them as "state science curriculum frameworks" or "science frameworks" for short.

received one of the 25 grants under the agency's Statewide Systemic Initiative (SSI) program (NSF, 1995). Also, the U.S. Department of Education—which had led in the funding of national standards for several school subjects—awarded three-year grants to 15 states and the District of Columbia to construct curriculum frameworks in science and mathematics. The frameworks were to serve as a bridge between national standards and the classroom. Both agencies thought the framework activity sufficiently important to accompany their awards to the states with funding for the evaluation of the resulting frameworks: NSF through a grant to CCSSO and the Department of Education through a contract to SRI International (1996). What follows draws largely on these evaluations as well as the author's experience in reviewing various state science frameworks on behalf of both evaluating organizations.

The Current Status of State Science Curriculum Frameworks

As of December 1994, 42 states had a state curriculum document covering science; 32 of these dealt with science only, 10 represented combined frameworks generally covering mathematics and sometimes other subjects, e.g., technology, as well (Blank & Pechman, 1995). At that time 25 states were developing new or revised curriculum frameworks in science; 16 had completed such frameworks between 1990 and 1994. A main purpose for frameworks completed since 1990 or currently undergoing revision was to provide high standards for students' science learning in the state. Teachers were identified as the primary users of the recent frameworks in the majority of cases;¹³ however, even the most detailed frameworks use a sufficiently large "grain size" to delineate science content, with just a few teaching examples scattered through the text, that most teachers would have difficulty using a state's framework for day-to-day instruction without further specification.

The more recent frameworks typically quote or refer to the national standards documents (most often *Benchmarks* (AAAS, 1992) or the most recent NRC draft available to them) as providing the basis for their own vision and rationale for science education. This evidences itself not only in content—one state goes so far as to adopt *Benchmarks* in toto—but also in style and pedagogic approach. For example, a much greater proportion of frameworks completed since 1990 use instructional examples to illustrate desirable classroom approaches than did earlier frameworks. Pedagogic strategies most frequently advocated include hands-on lessons and conducting experiments, using technology, and teaching such communication skills as writing about (or in) science and graphing. Many frameworks also discuss the need for assessments consonant with the content standards; only a few provide relevant examples, however, and many states continue to use tests that do not meet their own calls for assessments designed to further rather than impede the current reform efforts in science education.

Variation Among Frameworks

State documents examined by reviewers for CCSSO and SRI International varied in length from less than 20 pages to more than 500 pages, though—as with the national standards documents—some of these documents include supporting resources on assessment, implementation strategies, and professional development, while in other cases these are (or are planned as) separate state

¹³ Curriculum developers were cited as another primary audience nearly as often, in 75% of the frameworks contrasted to 80% citing teachers as primary audience.

documents to be used in conjunction with the framework. The CCSSO review (Blank & Pechman, 1995, p. 31) found that "content statements in frameworks vary from broad goals for mathematics or science content to specific topic and skill objectives for each K-12 grade." The report cites the following contrasting approaches in recent science frameworks (pp. 37, 38, 39):

Florida (1994): The Florida framework describes examples of concepts and activities by grade block within 8 broad knowledge strands of what students should be able to know and do, e.g., Strand 6: Processes of Life: (1) Design a model to demonstrate the correlation between healthful living and human body-system maintenance.

Ohio (1994): Ohio provides sample "grade level (instructional) objectives" in 4 strands of scientific knowledge, with about 50 objectives per grade, e.g., Strand, Scientific knowledge: Objective, Grade 7: The learner will investigate the interdependence and similarities of organisms and their environments (e.g., mimicry, camouflage).

Wyoming (1990): Wyoming defines broad goals for a science program, e.g., Components for Middle/Junior High Science: Exposing students to physical, earth and life science and relationships of these areas to their environment. Providing educational situations that incorporate laboratory, hands-on instruction, reading, discussion, research, and field experience.

The greatest commonality among the state frameworks is their organization of content along traditional disciplinary lines—life science, physical science, earth and space science—or some variant thereof. Not all follow this pattern, however. Massachusetts, for example, uses the following four content areas to illustrate the cross-cutting nature of science teaching and learning: Science as inquiry, Science as subject matter, Technological design, and Science and human affairs. Most of the frameworks organized along disciplinary lines also discuss cross-cutting themes; California's include Energy, Evolution, Patterns of change, Scale and structure, Stability, and Systems and interactions. Others stress process skills as cross-cutting learning to be accomplished. Generally, such areas as nature and history of science, science as inquiry, science applications, and science and society are stressed less in the state frameworks than in either the AAAS or the NRC documents. Also, although most state frameworks claim to follow the national standards documents, some have rewritten specific standards in such general language that they become unusable either for teaching or assessment purposes. An example cited in the SRI draft report (SRI International, 1996) illustrates the translation of one of the NRC standards providing specific knowledge to be learned about light, heat, electricity, and magnets (see p. 127 of *NSES*) into the following statement in a state science framework: "Identifying and describing the differences in the production and properties of light, heat, sound, electricity and magnetism."

With respect to illustrations of teaching practices in frameworks constructed since 1990, Massachusetts, Kentucky, and Nebraska are exemplary in offering contextualized examples or vignettes of a number of teaching strategies to demonstrate content implementation in the classroom. California, Florida, and New York also offer many examples of constructivist or activity-based lessons, assessment strategies, interdisciplinary connections, technological applications, and use of tools and educational technology.

One area of weakness in almost all the state frameworks is in dealing with equity issues. As do the national standards documents, most of the frameworks provide a vision or rationale for providing all students with full opportunity for science learning and expecting them to demonstrate high levels of achievement,¹⁴ but few follow through with suggestions on instructional approaches and activities aimed at reaching students currently less successful in science. Massachusetts is an outstanding exception; its framework provides not only a vision, but also teaching vignettes and strategies; assessment examples; discussions on the selection of materials, staff development, and teacher preparation; and recommendations for community involvement to further equity in science teaching and learning. Florida and Kentucky also deal with several of these areas, but the rest of the frameworks examined in the two evaluations either just set out the vision or provide examples in only one or two of the areas, most commonly regarding instructional strategies thought to be effective in reaching currently disaffected students.

Frameworks, Standards, and Reform

What is the relationship between a state science curriculum framework and the progress of science education reform in that state? Does an excellent framework guarantee effective reform? Clearly, the answer is no. If the framework is not built into the state's reform agenda, if it papers over broad policy disagreements, if it is not widely distributed and understood, if it is not accompanied by supporting resources and changes affecting curriculum, instructional materials, teaching practices, and assessment and testing, it is likely to be ineffective in driving improvement in science education, as the current assumptions underpinning systemic reform stipulate. Even when these conditions are met, political shifts may slow down or halt the implementation of the vision and goals embodied in the state's framework, as has been demonstrated in Arizona, California, New Jersey, New York, and Wisconsin.

But is a good state framework even necessary for science education reform, though it may not be sufficient? Do frameworks lead or follow reform activities? The authors of the SRI review (SRI International, 1996, p. 43) conclude:

Our evidence in the policy arena and in the classroom suggests . . . the frameworks projects were one of many reform activities under way in the states. Overall, the projects tend to reflect the reform efforts in the states rather than to lead them. This revised view of framework development and implementation rejects the mechanistic and linear imagery associated with the theory of systemic reform and replaces it with one that is more complex, nonlinear, and multifaceted.

Although this conclusion applies to state curriculum frameworks, it seems to this author equally applicable to standard setting activities at the national level. In view of all the activities AAAS, NRC, and NSTA are generating to assist implementation of the reforms urged in the standards documents they produced, they appear to share this opinion. As the director of the *NSES* project points out (Collins, 1995), standards documents, whether national or state-level, are policy

¹⁴ Alaska's draft Mathematics and Science Curriculum Framework provides a particularly strong equity statement commended by the reviewers of the 16 state frameworks developed with support from the Department of Education (SRI International, 1996, p. 31).

documents intended to influence and guide practice, but that is all they are. Ultimately, the changes these documents advocate must happen at the school and classroom level. For that to occur, the changes must be agreed to—in fact, "owned"—by parents, teachers, and administrators; enabled by the science knowledge and pedagogic skills of teachers; and supported by adequate resources, both material and organizational. This implies that the top-down, linear conception of "implementing" standards must yield to a strategy that combines approaches operating top-down (through policy influentials, scientists, and other experts), bottom-up (through teachers, parents, and principals), and through-the-middle (involving administrators and legislators at all levels). Above all, patience and persistence are imperative in continuing the reform effort in science education over the next decade—a conservative guess on the time needed before the reform's effects can be assessed.

References

- Atkin, J., Helms, J., Rosick, G. S., & Siner, S. (1996). Building on Strength: Science teaching in California public schools. In S. A. Raizen & E. D. Britton (Eds.), *Bold ventures: Case studies of U.S. innovations in science education* (Vol. 2). Boston: Kluwer.
- Aldridge, B. G. (1989). Essential changes in secondary science: Scope, sequence, and coordination. *NSTA Reports! 1*, 4-5.
- American Association for the Advancement of Science. (1989). *Project 2061: Science for all Americans*. Washington, DC: Author.
- American Association for the Advancement of Science. (1992). *Proposal to the National Science Foundation*. Unpublished.
- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford.
- American Association for the Advancement of Science. (1995). *Project 2061: Science literacy for a changing future, a decade of reform*. Washington, DC: Author.
- American Association for the Advancement of Science. (1996). *Science Education News*, 14(2).
- American Association for the Advancement of Science. (in press). *Designs for science literacy*. New York: Oxford University Press.
- American Federation of Teachers. (1995). *Making standards matter: A fifty-state progress report on efforts to raise academic standards*. Washington, DC: Author.
- Apple, M. (1992). Do the standards go far enough? Power, policy, and practice in mathematics education. *Journal for Research in Mathematics Education*, 5, 412-431.
- Atkin, J. (1994). Developing world-class education standards: Some conceptual and political dilemmas. In N. Cobb (Ed.), *The future of education: Perspectives on national standards in America* (pp. 61-84). New York: College Entrance Examination Board.
- Blank, R. K., & Epenshade, P. (1988). 50-state analysis of education policies on science and mathematics. *Education Evaluation and Policy Analysis*, 10(4), 315-324.
- Blank, R. K., & Pechman, E. M. (1995). *State curriculum frameworks in mathematics and science: How are they changing across the states?* Washington, DC: Council of Chief State School Officers.
- Bracey, G. (1993). The Third Bracey Report on the Condition of Public Education. *Phi Delta Kappan*, 75(2), 104-117.
- Britton, E. D., & Raizen, S. A. (Eds.). (1996). *Examining the examinations: An international comparison of science and mathematics examinations for college-bound students*. Boston: Kluwer.
- California Department of Education. (1985). *Mathematics framework for California public schools: Kindergarten through grade twelve*. Sacramento, CA: Author.
- California Department of Education. (1990). *Science framework for California public schools: Kindergarten through grade twelve*. Sacramento, CA: Author.
- Callahan, R. (1962). *Education and the cult of efficiency*. Chicago: University of Chicago Press.
- Catlin, S. (1986). New kids in class: Minorities pace growth in public schools. *California Journal*, 17(4), 189-192.
- Collins, A. (1995). National science education standards in the United States: A process and a product. *Studies in Science Education*, 26, 7-37.
- Crosswhite, F. J., Dossey, J. A., & Frye, S. M. (1989). NCTM standards for school mathematics: Visions for implementation. *Arithmetic Teacher*, 37 (3), 55-60.
- Donmoyer, R. (1995). The rhetoric and reality of systemic reform: A critique of the proposed National Science Education Standards. *COGNOSOS - The National Center for Science Teaching and Learning Research Quarterly*, 4(1).
- Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in science*. Philadelphia: Open University Press.
- Eisner, E. (1992). Curriculum ideologies. In P. W. Jackson (Ed.), *Handbook of research on curriculum: A project of the American Educational Research Association* (pp. 302-324). New York: Macmillan.

- Findings of the California study of elementary mathematics [Special issue]. (1990). *Educational Evaluation and Policy Analysis*, 12(3).
- Gagné, R. M. (1970). *The condition of learning* (2nd ed). New York: Holt, Rinehart & Winston.
- Harlen, W. (1985). *Teaching and learning primary science*. New York: Teachers College Press.
- Harlen, W. (1992). *Standards of practice for learner centered schools*. New York: National Center for Restructuring Schools and Learning.
- International Association for the Evaluation of Educational Achievement (IEA). (1988). *Science achievement in seventeen countries: A preliminary report*. Elmsford, NY: Pergamon.
- Lapointe, A. E., Mead, N. A., & Pillips, G. W. (1989). *A world of differences*. Princeton, NJ: Educational Testing Service.
- Lapointe, A. E., Askew, J. M., & Mead, N. A. (1992). *Learning science*. Princeton, NJ: Educational Testing Service.
- Marsh, D. D., & Odden, A. R. (1991). Implementation of the California mathematics and science frameworks. In R. A. Odden (Ed.), *Educational policy implementation* (pp. 219-240). Albany: State University of New York Press.
- McLeod, D. B., Stake, R. E., Schappelle, B. P., Mellissinas, M., & Gierl, M. J. (1996). Setting the standards: NCTM's role in the reform of mathematics education. In S. A. Raizen & E. D. Britton (Eds), *Bold ventures: Case studies of U.S. innovations in mathematics education* (Vol. 3, pp. 13-132). Boston: Kluwer.
- Mullis, I. V. S., Dossey, J. A., Foertsch, M. A., Jones, L. R., & Gentile, C. A. (1991). *Trends in academic progress: Achievement of U.S. students in science 1969-70 to 1990; mathematics, 1973-1990; reading, 1971-1990; and writing, 1984-1990*. Princeton, NJ: Educational Testing Service.
- Myers, M. (1994). Problems and issues facing the National Standards Project in English. In N. Cobb (Ed.), *The future of education* (pp. 259-276). New York: College Entrance Examination Board.
- National Assessment Governing Board. (1996). *Science framework for the 1996 National Assessment of Educational Progress. A report of the Council of Chief State School Officers with the National Center for Improving Science Education and the American Institutes for Research*. Washington, DC: Author.
- National Center for Education Statistics. (1992). *Digest of education statistics: 1992* (NCES 92-097). Washington, DC: U.S. Government Printing Office.
- National Center on Education and the Economy. (1995a). *Performance standards: Volume 1, elementary school*. (Consultation draft). Pittsburgh, PA: Author.
- National Center on Education and the Economy. (1995b). *Performance standards: Volume 2, middle school*. (Consultation draft). Pittsburgh, PA: Author.
- National Center on Education and the Economy. (1995c). *Performance standards: Volume 3, high school*. (Consultation draft). Pittsburgh, PA: Author.
- National Council of Teachers of Mathematics. (1980). *An agenda for action: Recommendations for school mathematics of the 1980s*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: Author.
- National Governors' Association. (1990). *National education goals*. Washington, DC: Author.
- National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Science Foundation, Division of Educational System Reform Programs. (1995). *The Systemic Reform "Report Card"* (draft). Arlington, VA: Author.

- National Science Foundation, Division of Research, Evaluation, and Communication, Directorate for Education and Human Resources. (1996). *Indicators of science and mathematics education 1995* (NSF 96-52). Arlington, VA: Author.
- National Science Teachers Association. (1992). *Volume 1: The content core, a guide for curriculum designers*. Washington, DC: Author.
- O'Day, J., & Smith, M. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 313-322). San Francisco: Jossey-Bass.
- Raizen, S. A. (1991). The reform of science education in the USA: Déjà vu or de novo? *Studies in Science Education*, 19, 1-41.
- Ravitch, D. (1995). *National standards in American education: A citizen's guide*. Washington, DC: Brookings.
- Robitaille, D., Schmidt, W., Raizen, S., McKnight, C., Britton, E., & Nicol, C. (1993). *Curriculum frameworks for mathematics and science*. Vancouver, Canada: Pacific Educational Press.
- Shavelson, R. J., Baxter, G. P., Pine, J., Yure, J., Goldman, S. R., & Smith, B. (1991). Alternative technologies for large scale science assessment: Instrument of education reform. *School Effectiveness and School Improvement*, 2(2), 97-114.
- Shymansky, J. A., Hedges, L. V., & Woodworth, G. (1990). A reassessment of the effects of inquiry-based science curricula of the 60s on student performance. *Journal of Research in Science Teaching*, 27(2), 127-144.
- Smith, M., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *Politics of curriculum and testing, The 1990 yearbook of the politics of education association* (pp. 233-267). London: Falmer.
- SRI International. (1996). *Evaluation of the Dwight D. Eisenhower mathematics and science education state curriculum frameworks projects: Second interim report*. Menlo Park, CA: Author.
- Tobin, K., Capie, W., & Bettencourt, A. (1988). Active teaching for higher cognitive learning in science. *International Journal of Science Education*, 10(1), 17-27.
- Welch, W. (1979). Twenty years of science curriculum development: A look back. In D. C. Berliner (Ed.), *Review of research in education* (pp. 282-306). Washington, DC: American Educational Research Association.