# CHAPTER 6
# BUILDING PROFESSIONAL CAPABILITY



The professional preparation of new teachers concentrates on getting them ready to teach certain subjects and grades in the existing curriculum, not on how to go about changing the curriculum they have inherited (let alone learning of the need to do so). Consequently, professional development of employed teachers tends to focus on improving their content background and instructional techniques. Similarly, the preparation of school administrators focuses largely on matters of school management, with little attention paid to the details of curriculum design.

If school districts are to achieve curriculum reform, therefore, it is essential that they build a professional capability for undertaking curriculum change. This chapter suggests how school districts can make headway in developing such a capability while at the same time beginning to make substantial improvements in some aspects of the curriculum itself. It calls upon educators in school districts to raise their collective level of science literacy, to become knowledgeable about the science, mathematics, and technology learning goals appropriate for all students, and to familiarize themselves with what is reliably known about student learning related to those goals. The chapter also calls for educators to increase their ability to make sound judgments about the quality of curriculum materials and to employ a variety of different curriculum formats that have particular instructional advantages.

Building professional capability for curriculum improvement in a school district calls for teachers to acquire certain knowledge and skills regarding science itself, how it is represented in literacy goals, how students learn challenging ideas, and how materials for instruction and assessment serve students' learning. It does not follow,

“A three-year study of education reform found that most staff development activities ‘were too short and lacked the follow-up necessary to develop the deep content and pedagogical knowledge necessary to meet new instructional goals…[and] did not appear to be building an infracturcture to promote and sustain teacher learning and instructional improvement over the long term.’”
—“The Bumpy Road to Education Reform” in *CPRE Policy Briefs* *(June 1996)*

“Curriculum-reform efforts are hard to sell and even more difficult to sustain if they can only succeed if teachers have special capacities, such as: extraordinary subject-matter expertise; the time, will, and skill required to develop their own curriculum materials; the ability to teach widely divergent students effectively; and the ability to maintain control over these students while allowing them freedom to learn on their own.”
—D. F. Labaree, “The Chronic Failure of Curriculum Reform” (1999)

however, that every teacher and administrator in a school district must attain the same level of expertise in every one of these matters. If one thinks of the faculty as a large team of professionals engaged in a shared endeavor, then it is the collective capability that counts—as long as there is sufficient collaboration among the members of the team. Teamwork is not common enough in the area of curriculum reform, but it is an aspect of professionalism that can be learned, if there is a will to do so. Therefore, the suggestions that follow are framed as team undertakings because the development of team skills is itself part of building professional capability.

## Increasing Faculty Science Literacy

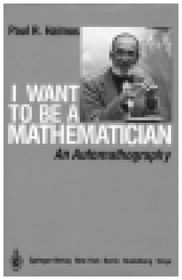
*Science for All Americans* argues that all high-school graduates should be science literate, and it describes the science, mathematics, and technology knowledge and skills that constitute such literacy. Although, in principle, all teachers should have reached that same level of science literacy, the present-day blunt truth is that too few of them—and too few college graduates in general—have done so. Perhaps someday all teachers will be science literate when they enter the profession, but in the meantime, steps need to be taken to enable teachers in a school district to make substantial progress toward achieving science literacy.

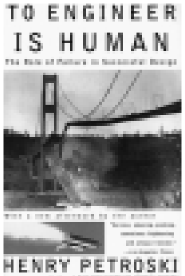
All of the ways suggested here require individual effort and administrative support. Learning takes time, resources, and encouragement. Without recognition by the community, school board, and administrators that teachers must upgrade their subject-matter knowledge continually, and without policy and budgetary support for the needed time, resources, and encouragement, in-service professional development will be of little consequence and contribute little to building districtwide professional capability. But with such recognition and support, teachers can improve their understanding of science, mathematics, and technology by engaging in individual or group study of selected readings or growth-of-understanding maps, setting up a long-term program of workshops, or by taking appropriate courses.

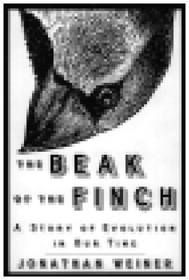
### Readings

Because it defines adult science literacy, *Science for All Americans* can provide a focus for faculty study. It does not serve well as a textbook, its purpose being to summarize rather than to teach, but excellent books and articles are available that cover in detail most of the topics in *Science for All Americans*. Project 2061 has undertaken the task of sorting through the tens of thousands of books marketed to the general public that deal with science, mathematics, and technology to find those that are the best for this aspect of professional development. *Resources for Science Literacy: Professional Development* is a CD-ROM containing five different kinds of resources linked to *Science for All Americans* (see next page). Among them is a compendium of what Project 2061 believes are some of the best books available for individual and group study. Each book is described, published reviews of it are reproduced, and its ties to *Science for All Americans* are specified.

Descriptions of recommended trade books can be found on *Designs on Disk* and on Project 2061's Web site at www.project 2061.org.


Although teachers can use the CD-ROM to design and pursue individual programs of study, group study should be encouraged. One approach is to form reading groups in each school. After agreeing on a topic, each group discusses the reading possibilities and then selects a book for everyone to read and discuss in subsequent meetings.

Study groups will differ in the number of participants, whether teachers from other schools and parents are invited to attend, how many books are taken on each semester, how sessions are conducted, and so on. Within reasonable limits, such variations are not likely to affect the outcome greatly. What is important is that the readings be selected by the group itself and that participating faculty can earn appropriate professional development credit. It is desirable, of course, that time for the group to meet be included in the formal school schedule. But the fact that scheduling practices in most schools often make it difficult for teachers to meet together during the school day need not be an impenetrable barrier. Each group should be able to find one evening, late afternoon, or early morning once a month on which to meet to discuss the reading. Alternatively, electronic conferencing can make it possible for members of a groups to participate in the conversation at their convenience. (Indeed, the group can use the World Wide Web to look for reviews and expert commentary on the book under discussion.)

*Designs on Disk* contains a database with convenient forms for recording reactions (positive and negative) to each book a study group reads. This process also enables the group to keep track of topics that have been studied by at least some members of the faculty and to identify teachers who are well informed on particular aspects of science, mathematics, and technology and can be called on as consultants by other teachers.

A revised edition of *Resources for Science Literacy: Professional Development* will also include a database of newspaper, journal, and magazine articles that shed light on topics that are central to science literacy.

## CONTENTS OF *RESOURCES FOR SCIENCE LITERACY: PROFESSIONAL DEVELOPMENT*

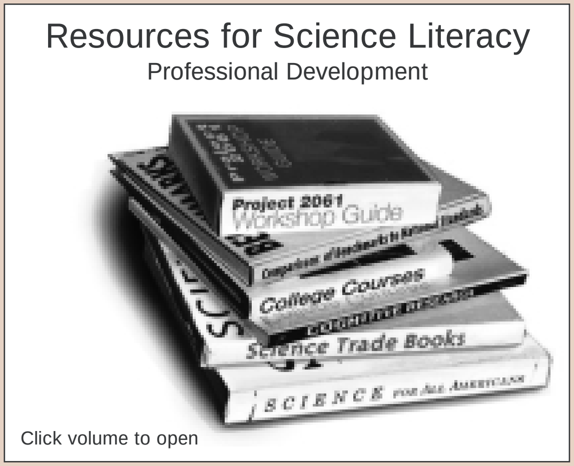
SCIENCE FOR ALL AMERICANS
The full text of Project 2061's landmark report is available for the first time in an electronic format. Links to other components allow users to identify resources on the CD-ROM that are relevant to specific chapters and sections of *Science for All Americans*.

PROJECT 2061 WORKSHOP GUIDE
The Workshop Guide contains a variety of presentations, scripts, activities, and supplementary materials that can be used to design and conduct Project 2061 workshops.

COMPARISON OF *BENCHMARKS FOR SCIENCE LITERACY* TO NATIONAL STANDARDS
Detailed analyses compare *Benchmarks* to national content standards developed by the National Research Council, the National Council of Teachers of Mathematics, and the National Council for the Social Studies.

COGNITIVE RESEARCH
An introduction to current cognitive research literature, along with *Benchmarks* Chapter 15: The Resarch Base and its accompanying bibliography of more than 300 references to the educational research literature, sheds light on how students learn particular concepts from *Science for All Americans* and *Benchmarks*.

This graphic menu from *Resources for Science Literacy: Professional Development* displays the contents of the CD-ROM.




SCIENCE TRADE BOOKS
Full bibliographies, reviews, and other descriptive data are provided for more than 120 books for general readers dealing with many areas of science, technology, and mathematics. Each book is linked to related *Science for All Americans* chapters and sections.

COLLEGE COURSES
Descriptions of 15 undergraduate courses suggest how to teach college students particular concepts from *Science for All Americans*. The syllabi are linked to relevant chapters and sections of *Science for All Americans*.

**Project 2061 Strand Maps**

Study of these maps, which attempt to depict how students' understanding might develop over the school years, is a useful adjunct to a program of readings. The strand map example on the next page shows how the development of a concept can be traced from its simple beginnings as ideas join and grow in sophistication. Study groups may find it useful to work their way up a map, discussing its individual benchmarks and seeking information from the books on the reading list. Although the main purpose in studying such maps is usually to acquire a better understanding of the development of student learning or to plan curriculum sequences, many teachers have found that the process serves as an excellent organizing device for helping them improve their understanding of the topics—and to decide what ideas to teach and what to emphasize about them.

**Courses**

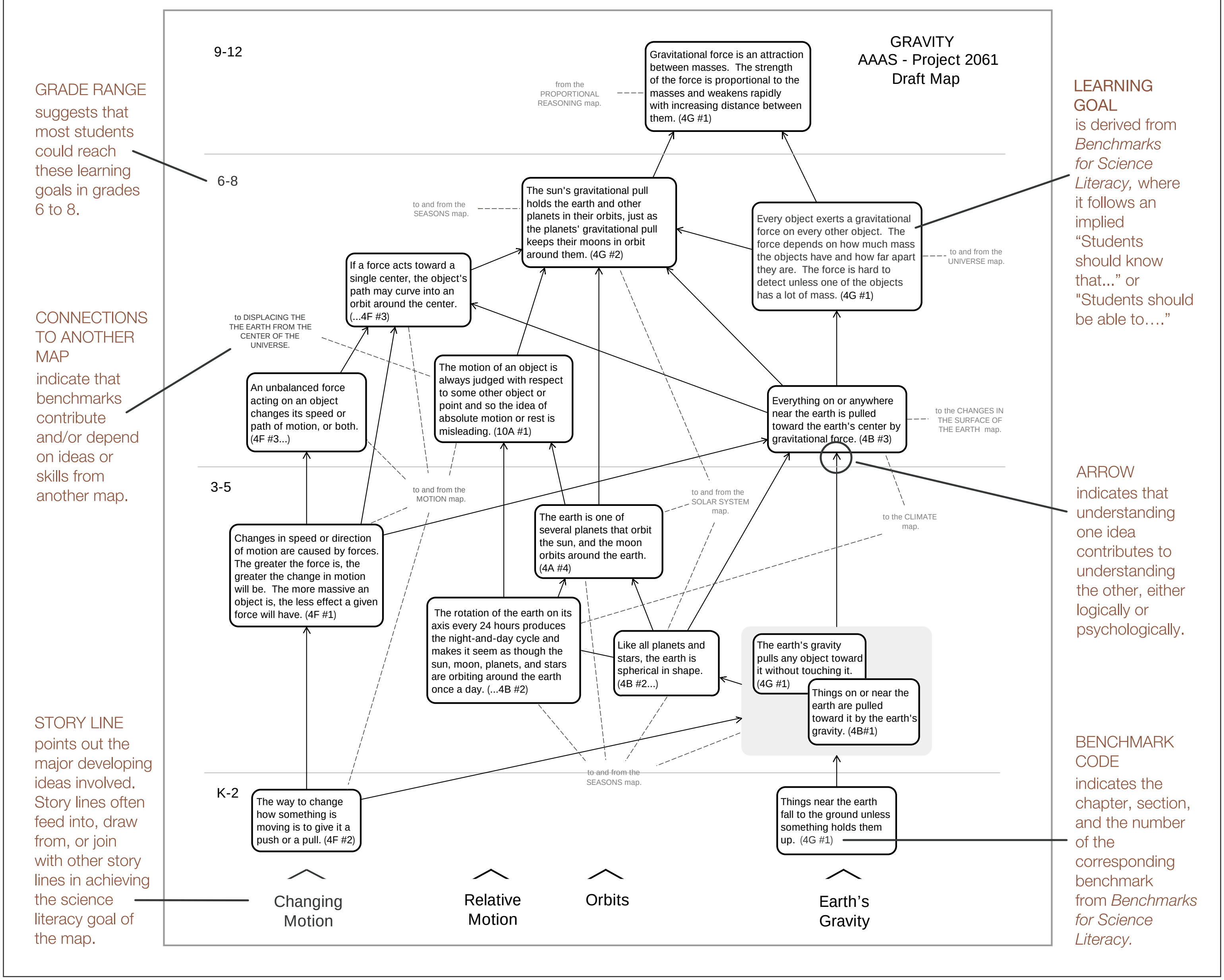
A time-honored way for teachers to acquire content knowledge and develop professional skills is to take college courses in science, mathematics, and technology. Many colleges and universities make such courses available on campus during the summer and academic year, but the institutions may not be within geographical or financial reach of teachers. And even if within reach, they may not offer the content courses that the teachers need to make progress toward science literacy. Once a teacher has reached a comfortable level of literacy in a particular area, regular college courses may be useful. The priority, however, should be on achieving literacy efficiently.

Colleges and universities are becoming more willing to tailor the content, instructional style, and scheduling of courses to fit the specific needs of groups of teachers. As a result, taking courses may become an important component of a school district's overall plan for building its professional capability for curricular reform. The concomitant increase in video and/or computer-based courses available at a distance can enhance that effort. Some universities are developing courses that will serve the science literacy needs of teachers, both preservice and in-service, by focusing on the image of literacy portrayed in *Science for All Americans* and accommodating the research on the prevalence of misconceptions in many areas. Such courses can improve not only the teachers' grasp of science, mathematics, and technology, but how those subjects can be taught effectively.

The use of some combination of reading groups focusing on content, study sessions built around strand maps, and courses and minicourses (summer and academic-year,

The syllabi of some tailored courses can be found on the *Resources for Science Literacy: Professional Development* CD-ROM, and more examples will be added in subsequent versions. Perhaps these syllabi will motivate many other colleges and universities to contribute to the database—especially if there is a clear demand for such courses from groups of teachers in a school district.

**A PROJECT 2061 STRAND MAP**



direct and by way of the Internet) can gradually lead to an increase in the proportion of a school district's faculty members who are science literate (as spelled out in *Science for All Americans*). With careful planning and some luck, summer research opportunities offered by local business and industry can contribute as well. In many locations, national laboratories also provide programs for teachers. Because of the years of service that lie ahead of younger teachers, and the need for them to have the habit of continuing education become ingrained, it is particularly important that they be included in this professional-development process. Simultaneously, school districts should raise their hiring standards for new teachers by stating explicitly that evidence of science literacy will be taken into account when hiring and by notifying the relevant teacher-education institutions of their expectation that candidates be science literate.

## Understanding Student Learning Goals

Having a faculty that is well grounded in science, mathematics, and technology is not enough to ensure that all (or even most) students will learn enough in school to become science literate. Research studies have shown that teachers' subject knowledge is only part of the story of successful learning. Equally important is their understanding of precisely what it is that they expect students to learn, the developmental pace at which students are able to learn those things, and the difficulties that students typically encounter.

Also under way: *Technology for All Americans,* a report of the International Technology Education Association, will spell out learning goals in technology education.

Fortunately, a faculty does not have to determine for itself what appropriate student learning goals are for science literacy. The efforts of the nation's scientific and science teaching organizations over a period of years have resulted in publication of *Benchmarks for Science Literacy*, *Curriculum and Evaluation Standards for School Mathematics*, and *National Science Education Standards* (*NSES*). These reference works are in general accord on the importance of reducing the mass of an overstuffed curriculum and specifically on what science, mathematics, and technology knowledge and skills are most important for students to learn. Even if they are empowered to create their own science-literacy learning goals, local groups will find it valuable to study these reference works in detail (as distinct from just comparing topic headings).

None of the national groups has merely made a selection of assorted topics. They have attempted to identify interconnected sets of ideas and skills that will, in the *Science for All Americans* phrase, "maximize students' ability to make sense of the world and to learn more about it." Reformers should take care not to disregard the coherent set of specific learning goals in the national documents or to simply pick and choose

For a more detailed discussion of the term "topics," see CHAPTER 7: UNBURDENING THE CURRICULUM.

casually among them. By so doing, they may lose not only important interconnections within or across topics, but also the potential for K-12 continuity that helps students to gradually build their understanding of difficult concepts. Reformers should also beware of simply *adding* national-goal recommendations to the requirements of an already unwieldy curriculum. The national goals for science literacy are designed to help educators focus on fewer, but more important, ideas so that all students have a chance to learn them well.

Understanding the real intent of a set of specific learning goals is not as straightforward as it may seem at first glance. The difficulty comes from taking the benchmarks and standards to be lists of "topic headings" (as is often the case with familiar curriculum guidelines), rather than as painstaking selections and specifications of the essential aspects of ideas to be learned and understood in relation to one another. For example, seeing the section heading "Cells" in *Benchmarks* could be taken as an endorsement to teach anything whatever about cells—including over a hundred technical terms typically found under the topic of cells in the high-school biology course—rather than the carefully chosen ideas that *Benchmarks* describes.

"Topics" have another muddling effect besides excessive inclusiveness. They often identify what is to be studied, without specifying just what is to be learned. As noted in Chapter 4: Curriculum Blocks, "acid rain" is a likely topic for a middle-school science unit. Neither *Benchmarks* nor *NSES* includes acid rain as a high-priority component of science literacy. Nonetheless, studying the topic of acid rain could help students toward any number of benchmarks that are high-priority components having to do with differences in climate, the mechanics of the water cycle, the appropriateness of measurements, fitting data with mathematical models, proportionality of concentration, the difficulty of anticipating side effects of technology, uneven benefits and costs of trade-offs, and so on. From the perspective of specific learning goals, acid rain is a *context* in which many such benchmarks can be pursued. The crucial distinction between what is to be learned (specific learning goals) and what is to be taught (topical context) is often lost in education discourse, with the former being taken erroneously to be synonymous with the latter. It is essential to keep the distinction straight.

How, then, can a school district foster the needed understanding of student learning goals among its faculty? One practical approach, framed here with reference to the science literacy learning goals set out in *Benchmarks,* is to use Project 2061 tools: to study the growth-of-understanding maps; analyze instructional topics against specific learning goals; and participate in the kinds of  the workshops described on

*Resources for Science Literacy: Professional Development*. The next three sections describe these three activities.

**Studying Strand Maps**

Teachers report that the Project 2061 maps about students' growth of understanding are especially valuable when used by small study groups of elementary-, middle-, and high-school teachers all working together. Larger groups can form small subgroups and have them compare their interpretations from time to time. A common way to proceed is to follow three basic steps:

First, the group should start with maps on familiar topics that most of the group members feel comfortable with, and then move on to those perceived to be more complicated or less familiar as the group's confidence builds. "The Water Cycle," "Culture and Heredity," and "The Conservation of Matter" are generally well received by teachers undertaking the study of growth-of-understanding maps for the first time.

Second, the group should generally work from grade K-2 benchmarks toward ones for grade 9-12, although there is likely to be a lot of back and forth. Considering one benchmark at a time, the group members should first discuss what they think it means and doesn't mean. They should read the appropriate section of *Science for All Americans*, then the essay in the section of *Benchmarks* in which the benchmark under study is found, and then the relevant research findings, if any, cited in Chapter 15 of *Benchmarks*. Then the group should once again discuss the benchmark.

Third, the group should run through the map again, this time discussing the connections among some of its benchmarks. The group members should consider what each arrow means—whether it indicates a necessary prerequisite or only a helpful contribution. Seeing the relationships between two or three benchmarks will help to clarify the meaning of each of them and to make explicit what is meant by growth of understanding.

*Benchmarks on Disk, Resources for Science Literacy: Professional Development,* and the Project 2061 Web site include several sample maps and commentary on them. Many more are scheduled to appear in *Atlas of Science Literacy*. Further suggestions on using maps appear in CHAPTER 8: INCREASING CURRICULUM COHERENCE.

**Analyzing Instructional Topics**

Faculty study groups should analyze instructional topics in the light of specified learning goals. Briefly, there are three steps to be taken:

First, the group should identify a contextual topic (for example, "Cloning") that it believes offers a rich opportunity for learning, and then try to agree on which benchmarks, if any, could be targeted in studying that topic at each of the indicated grades.

Second, referring to adopted textbooks, course outlines, and curriculum frameworks, the group should check the list of specific learning goals (benchmarks) against

### CLARIFYING A PARTICULAR BENCHMARK

In Project 2061 workshops, the full meaning of a benchmark is revealed through careful consideration and study of the following:

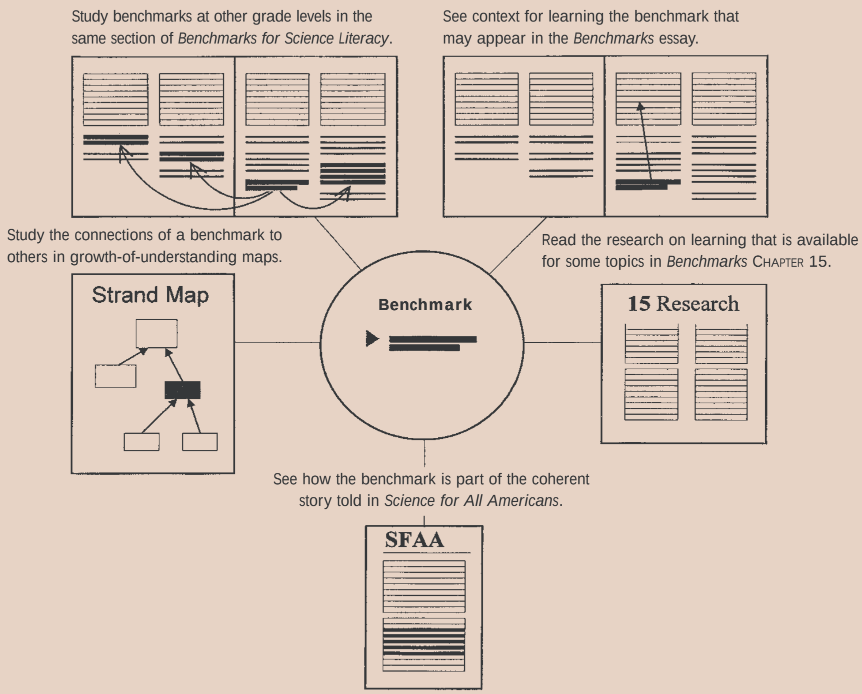
- **Adult Literacy Goal.** For each *Benchmarks for Science Literacy* section, there is a corresponding *Science for All Americans* section describing adult science literacy goals for that topic; it can help participants understand where benchmarks in that section are aiming.
- **K-12 Context.** A review of benchmarks for other grade levels from the same *Benchmarks* section helps participants understand the level of sophistication intended by the benchmark.
- **Instructional Strategy.** The introductory essays in the *Benchmarks* section for the benchmark being studied help participants understand difficulties students may have with the concept or skill and offer some suggestions for helping students achieve the benchmark.
- **Research Base.** Summaries of research on the topic from *Benchmarks* Chapter 15: The Research Base suggest likely limitations in student understanding of the benchmark and, therefore, imply an appropriate grade level for the benchmark. They also point participants to the original research articles.
- **Strand Maps.** A relevant conceptual strand map depicts K-12 growth of understanding for a particular *Science for All Americans* idea. The maps on which a benchmark appears help participants see how other benchmarks relate to and contribute meaning to the benchmark being studied.

There is no special order to these activities, although elementary teachers often prefer to see where an idea leads next in strand maps, whereas high school teachers often look to the eventual context of understanding in *Science for All Americans*.

Consider the grade 6-8 benchmark on the flow of matter and energy: “Food provides molecules that serve as fuel and building materials for all organisms….” Educators who read the precursor grade 3-5 benchmark, “Almost all kinds of animals' food can be traced back to plants,” are more likely than those who don't to realize that the grade 6-8 benchmark goes beyond “what eats what.” By reading the essay, educators learn that following matter through ecosystems needs to be linked to study of atoms, which itself is impractical before late middle school. Research alerts educators to the difficulty students have accepting the idea that plants make their only food from water and air. And the strand map Flow of Matter and Energy shows how understanding the conservation of matter in living systems depends on understanding the structure of matter.

what is actually taught. The group's task here is to spot where mismatches seem to occur—such as aspects of the topic being taught that are unrelated to any of the selected benchmarks, or instruction having no bearing on some of the benchmarks, or instruction relating to some of the benchmarks being positioned earlier or later than those benchmarks. The members of the group should review the selected benchmarks to make sure they understand them, using the techniques described in the box on the facing page and the one below. They can then modify the topic/goal list to obtain a statement of learning goals for the topic that teachers at all grade levels can support.

## USING PROJECT 2061 TOOLS TO CLARIFY A BENCHMARK




Third, the group should continue with new topics, gradually building a record of what topics and goals will be part of the curriculum at each grade level and what changes have to be made in the curriculum to accommodate the decisions. As this process continues, the participating faculty members will be able to increase their knowledge of student learning goals and hence help to raise the capability of the faculty as a whole to engage in curriculum reform.

### Conducting Benchmark Workshops

The CD-ROM *Resources for Science Literacy: Professional Development* describes workshops designed to help educators understand *Benchmarks* (and, by virtue of their similarity, *NSES* and NCTM's *Standards* as well). Its many embedded options enable a person or group to create a workshop of almost any duration tailored to the particular interests, needs, and circumstances of a group of educators. As the summary of the CD-ROM's workshop component opposite suggests, workshop designers have access to outlines, scenarios, handouts, transparency masters, references, and sample presentation scripts to produce a wide variety of workshops for their colleagues.

## Becoming Familiar with Research on Learning

Teaching is more craft than science. Over the years, teaching has been shaped informally by what seems to work best in the individual teacher's classroom, by broad theories about student learning, and by tradition. But for the most part, practices based on those experiences, theories, and traditions have not been put to rigorous tests to see if they really are sound. Similarly, teachers and other educators (textbook developers, framework writers, test makers) necessarily have had to base curriculum decisions on shared beliefs that have accumulated over the years on what students can learn, beliefs that until recent years have rarely been examined systematically.

Gradually, however, research is identifying some principles of teaching and learning that apply rather generally and some that apply to specific content areas. For example, giving a student enough time to formulate an answer thoughtfully (the well-known "wait time") may be a universal principle of good teaching. Much more specifically, quite a bit is known about when and how students seem to be able to learn about (and believe in) the molecular model of matter, the shape of the earth, the distinction between heat and temperature, how plants make their food, natural selection, and the equivalence of fractions and decimals. Currently, there are still large numbers of concepts for which little learning research has been carried out. Perhaps the focus offered by the reduced and highly specific ideas in national goals will stimulate interest in—and funding to support—a greatly intensified research effort on how to promote learning of those ideas.

There are many accounts of creative teaching ideas to be found in teacher magazines and journals, which usually focus on whether lessons "work" in the classroom. This literature can be very stimulating, but seldom offers convincing evidence for what particular ideas or skills students may have learned. The advent of specific national learning goals may help to focus that literature, too, on demonstrating results.

## DESIGNING PROJECT 2061 PROFESSIONAL DEVELOPMENT WORKSHOPS

The Workshop Guide, found in the Project 2061 CD-ROM *Resources for Science Literacy: Professional Development*, provides advice, example scripts, and materials for designing a variety of workshops—including a large store of transparencies, handouts, and readings.

All Project 2061 workshops aim to demonstrate the need for reform and then a particular way in which Project 2061 tools can help (for example, analyzing curriculum frameworks). Workshops usually have three major stages: (1) general reform rationale, (2) particular use of tools, and (3) reflection and summary. Within each stage, two to ten options are provided for each of several steps. After first deciding which tool use to focus on, workshop designers select other options that will best suit it, meet the participants' needs, and fit the time available. To help prospective workshop leaders get started, three examples of complete workshop agendas are also provided for six-hour, one-day, and two-day workshops. The Workshop Guide utility itself can also serve as a tutorial for learning more about Project 2061 and how to use its tools for science literacy.

This is the general workshop format as presented in *Resources for Science Literacy*:

### Opening

This stage allows the workshop leader to find out more about what participants already know about science education reform and Project 2061 and to establish the specific learning goals for the workshop.

- Introduction (four options available)
- Need for Change in Science Education (six options available)
- Workshop Goals (four options available)
- What Do Participants Know about Project 2061? (three options available)

### Project 2061 Tools

This is the core of a Project 2061 workshop, and the option chosen here will help determine which Opening and Closing options are selected.

- Overview of Tools Available from Project 2061 (ten options available)
- Exploring the Use of Project 2061 Tools to (choose one of the following uses):
  - Understand the nature of benchmarks (five options available)
  - Analyze curriculum frameworks (one option available)
  - Analyze curriculum materials (two options available)
  - Analyze instruction (five options available)
  - Design instruction (seven options available)

### Closing

In this stage workshop participants reflect on what they have learned and provide the workshop leader with feedback on the effectiveness of the workshop itself.

- Summary (five options available)
- Evaluation (six options available)

As the body of systematic knowledge about teaching and learning grows, educators can turn to it more frequently than in the past for guidance in making informed curriculum decisions. This does not mean, however, that all educators in a school district need to be trained as researchers, or even as skilled analyzers of research. The professional capacity of a school system can be enhanced if some of the faculty accept responsibility for keeping up to date on what the research (see marginal note) says about learning and teaching and for locating pertinent research information when needed. Several different teams should be formed to track research for the school district, since most cognitive research on learning focuses on the learning of particular concepts of a particular subject at a particular grade level. Distribution of interesting findings to faculty in the relevant areas would, of course, be part of this task.

A school-district group concerned with research on learning in science, mathematics, and technology could start by becoming acquainted with the scope of cognitive research in those fields and with what implications it has for practice. From time to time, members of the group will want to read and discuss some of the original research accounts to see precisely what was done and found out. But most often their task will be to relate research findings to curricular issues.

"Even after some years of physics instruction, students do not distinguish well between heat and temperature when they explain thermal phenomena (Kesidou & Duit, 1993; Tiberghien, 1983; Wiser, 1988). Their belief that temperature is the measure of heat is particularly resistant to change. Long-term teaching interventions are required for upper middle-school students to start differentiating between heat and temperature (Linn & Songer, 1991)."
—*Benchmarks for Science Literacy,* p. 337

From *Benchmarks* Chapter 15: The Research Base, the group could turn to the more elaborate, annotated references to the research literature and to the summary accounts of research provided in *Resources for Science Literacy: Professional Development*, which are also keyed to sections of *Science for All Americans* and *Benchmarks*. Other helpful places to look for findings from research are the *Handbook of Research on Science Teaching and Learning* (Gabel, 1994) and the corresponding *Handbook of Research on Mathematics Teaching and Learning* (Grouws, 1992); the series of volumes entitled *What Research Says to the Science Teacher* and the mathematics collection *Research Ideas for the Classroom* (Jensen, 1993; Owens, 1993; Wilson, 1993) with volumes for three different grade ranges; and the special mathematics, science, and technology chapters in *Handbook of Research on Curriculum* (Jackson, 1992).

## Learning to Analyze Curriculum Materials

Educators select instructional materials to serve a curriculum, or perhaps it is the other way around—materials that educators select determine the curriculum. Either way, making decisions about instructional materials is a major professional responsibility in every school district, even in ordinary circumstances. When a district is engaged in curriculum reform, the evaluation of instructional materials takes on still greater importance. Building a professional capability for reform therefore includes making sure that those individuals who will make decisions about the selection of curriculum materials acquire the technical knowledge and skills for analyzing materials and comparing their advantages and disadvantages. Moreover, teachers who are not themselves directly engaged in the process should know enough about it to be able to interpret and respond to the recommendations.

Asking whether curriculum materials would actually help students to achieve benchmarks is a powerful way to understand both better. In the past, the evaluation of curriculum materials has often been sporadic (taking place for any subject only once every five years or so), free floating (paying little or no attention to articulation across the grades), ad hoc (being conducted each time by a new committee of mostly novice evaluators), and short (having to be accomplished usually in a matter of hours or days on top of regular assignments). The evaluation process typically has been highly subjective, with little basis for estimating whether the conclusions reached are consistent ("reliable") and true ("valid").

Judgments by independent reviewers are said technically to be "reliable" if they are similar to one another; whether they are true is another matter. Reviewers may share a bias or misunderstanding, and hence make consistent but false judgments.

Shortcomings of time and effort are matters of administrative priority and can be easily changed once the task of curriculum-materials analysis is taken seriously enough. More difficult to correct is the absence of two resources: (1) a set of coherent and authoritative specific learning goals as the chief criterion for judging curriculum materials, and (2) a rigorous analytical procedure for examining curriculum materials in the light of their likely contribution to the achievement of those learning goals.

As energetic and conscientious as the efforts of some educators have been to design sound curriculum without these resources, a new level of effectiveness is now widely possible—at least in science, mathematics, and technology education. National benchmarks and standards provide the needed learning goals at a suitable level of specificity. And even as Project 2061 was developing *Benchmarks*, it began exploring how specified learning goals could actually be used effectively to help make decisions about such practical matters as teaching, testing, and curriculum materials. A system of analysis has been developed by Project 2061 that focuses on the attainment of specific learning goals and describes the kind of evidence required to make a case for what students are likely to learn. See the following pages for a description of the system and the criteria it uses.

### ANALYZING A CURRICULUM MATERIAL

**Step 1: Identify likely benchmarks.** Make a list of a few benchmarks that are important and that you would expect the material to focus on. Next, look through the material to find instructional experiences that might help students learn those benchmark ideas. If you can't find such evidence for a particular benchmark, then cross it off your list. That will give you a much shorter list of benchmarks on which the material actually focuses.

**Step 2: Clarify the benchmark's meaning.** Pick one benchmark from the short list and study it, as described on pages 188 and 189.

**Step 3: Reconsider how explicitly the material targets the benchmark.** Now go back and briefly describe the evidence and where you found it in the material, including the teacher guide. Consider whether the activities are appropriate for the intended grade level. You may find, for example, that a set of lessons on the water cycle, although advertised as K-2, focuses on evaporation and hence targets benchmarks for grades 3-5. For activities that are appropriate, is there adequate guidance to ensure that the benchmark idea will be addressed?

**Step 4: Estimate how effective the instruction would be.** Now use the best available knowledge of how students learn to reflect on how much students would actually learn about the benchmark from the recommended instruction. Project 2061 has developed criteria for estimating the effectiveness of instruction. (See facing page.) For example, do the recommended activities provide students with *memorable* experiences, opportunities to *reflect* on them, and opportunities to explore concepts in *varied contexts*? Even good teachers find that they often underestimate how difficult some ideas are for many students, especially when the students already have persistent misconceptions.

**Step 5: Summarize and make recommendations.** Review your findings on all the criteria and summarize how effective you think the material, together with what *appears explicitly* in its teacher guide, would be for helping students to achieve the benchmarks. Consider what it would take to improve it—for example, increasing the variety of phenomena studied, providing better questions to guide students' thinking, and making the assessment fit the specific benchmarks better. Finally, recommend how the material should (or should not) be used in the curriculum.

## CRITERIA FOR ESTIMATING INSTRUCTIONAL EFFECTIVENESS

Project 2061's curriculum-materials analysis procedure uses the following questions to determine the extent to which a material's instructional strategy is likely to help students learn the content. Each question focuses on *specific benchmarks*, not just content in general. To what extent does the material:

**Provide a sense of purpose? That is, does it**
… provide an overall sense of direction that students will understand and find motivating?
… provide a sense of purpose for each lesson and its relationship to others?
… provide an obvious rationale for the sequence of activities (versus just a collection)?

**Take account of student ideas? That is, does it**
…specify prerequisite knowledge/skills?
…alert teachers to commonly held student misconceptions?
…suggest how to find out what students think about relevant phenomena and principles?
…explicitly address commonly held student ideas?
…include suggestions for teachers on how to address ideas that their students hold?

**Engage students with phenomena? That is, does it**
…include direct experiences (or close approximations) with relevant phenomena?
…promote experiences in multiple, different contexts to foster generalizations?

**Develop and use scientific ideas? That is, does it**
…build a case for scientific ideas based on their success in explaining phenomena?
…link technical terms to experiences and only when needed for communication?
…include a variety of representations of scientific ideas that are both comprehensible and valid?
…tie ideas together over time logically and explicitly?
…explicitly draw attention to appropriate connections among benchmark ideas in different topics?
…describe how teachers can demonstrate application of skills or knowledge?
…provide tasks on which students can practice application in a variety of situations?

**Promote student thinking? That is, does it**
…encourage each student to express, clarify, and justify—and get feedback on—his or her ideas?
…include sequenced tasks or questions to guide student reasoning?
…help or suggest how to help students to know when to use knowledge and skills in new situations?
…suggest how students can check their progress and consider how their ideas have changed?

**Assess progress? That is, does it**
…include tasks to assess student achievement of particular benchmarks?
…include tasks requiring new application, not just plugging into formulas or repeating definitions?
…embed assessment in instruction with advice on using the results to choose or modify activities?

**Promote other benefits? That is, does it**
…improve teachers' own understanding of science, mathematics, and technology and their connections?
…foster student curiosity, creativity, and healthy questioning?
…foster high expectations and success for all students?
…explicitly draw attention to appropriate connections to ideas in other units?
…have other features worth noting?

The CD-ROM *Resources for Science Literacy: Curriculum Materials Evaluation* has detailed explanations of these indicators and provides both good and poor examples for each.

Even when highly able teachers are carrying out the evaluation procedure, it is good practice for them to follow detailed steps aimed at counteracting old short-cut habits that were acquired under typical conditions of inadequate study time. Separate, summary judgments about content match and instructional quality are often seen not to hold up when the particular instruction for particular ideas is investigated. The key question has to be, "How likely is it that all students will learn this particular idea from these prescribed activities?" There is no question that the analysis procedure is demanding of time and effort, but experience with a variety of simpler alternatives shows that reliable and valid judgments require such investment. It is possible however, that experienced analysts will eventually be able to shorten the procedure.

The Project 2061 procedure for analyzing curriculum materials needs to be conducted by teams that already have expertise in content and instruction and are able to use the analysis criteria accurately and reliably—which requires intensive training and practice. That is a tall order, and many school districts may well find it too daunting to undertake. For that reason, Project 2061 is in the process of training national teams to evaluate curriculum materials and is slowly building an on-line database of science, mathematics, and technology materials that have been evaluated by those teams. The teams' evaluation reports will not make yes-or-no recommendations but rather present profiles showing how well the materials appear to support student progress toward benchmarks (or other coherent sets of specific learning goals). But even when such analyses become available, they are not likely to be used well unless users understand enough of the process involved to make sense of the profiles and to determine their implications for making decisions. *Resources for Science Literacy: Curriculum Materials Evaluation* provides guidance for developing both general familiarity with the process and the technical skills required for its implementation.

## Acquiring Curriculum Versatility

Project 2061 does not advocate any particular kind of instructional format, scheduling of time, or form of assessment. Nor does it endorse any particular organization of instruction—thematic or academic, integrated or disciplinary, cooperative-group or competitive, lecture or hands-on. Different content and learning goals are likely to be better served by some formats than by others. Moreover, a wide variety of research findings suggest that different students learn different ideas in different ways. But the evidence does not yet provide a sound basis for predicting the best way for a particular student to learn a particular idea from a particular teacher. Perusal of the literature on trials of seminars, design projects, independent study, etc. show that these have seldom had the benefit of well-defined goals and well-controlled circumstances. Again, the improved focus afforded by national goals may expand and improve the knowledge base. We can be fairly sure that no single approach to the organization of instruction will be found to be consistently best for all students. So in anticipation of documented advantages and disadvantages of different instructional strategies, it is a good idea to develop some collective experience with a variety of formats.

Future developments in curriculum design and in instructional technology may make possible many more options for organizing instruction than are now available. High-quality curriculum blocks developed in the future may require a variety of different instructional strategies. For a school district to consider those possibilities seriously, it needs to appreciate their individual advantages and limitations and know what it takes to operate them well with local students. Project 2061 is not pressing for maximizing curriculum variety as an end in itself, but for freeing districts to use the latest resources available, regardless of their format. If none of the faculty in a school district has firsthand experience with other than the traditional way of operating a curriculum, then there will be no basis for making informed decisions when alternatives are proposed. To develop the needed professional capability, a school district should identify groups of teachers who will, through study and experience, become knowledgeable in the conduct of different instructional formats and time patterns. In addition, there ought to be teachers at every grade level who have used the various formats in actual instruction.

## Alternative Instructional Formats

No one teacher needs to be expert in the conduct of every possible format, but collectively a school district should have teachers with experience in the use of a variety of formats. In this way, a district will have a base for making informed decisions on the adoption of different formats—and to be in a strong position to introduce new formats if creditable curriculum blocks require them.

A convenient way to get started building such a capability is to set up a conference “listserv” on the district computer network for each of the formats of interest at each

Although there is only sparse research on how well alternative formats work, *Designs on Disk* includes a variety of articles about developing and using them.

The *Designs on Disk* database for reporting on and discussing instructional formats offers a convenient source of ideas for school districts to test on a limited scale.

grade level. Thus, there could be elementary-school, middle-school, and high-school conferences for teachers and administrators interested in seminar, project, independent-study, peer-teaching, integrated-course, and discipline-based course formats. The members of each conference would share in building an electronic library of outstanding articles having to do with the theory and application of its format. A next step would be for some teachers to implement one or more of the proposed ideas. That would entail describing the process, keeping notes on what transpired, and sharing the results with the conference members.

Consider the following. An elementary-school teacher may decide to try both independent study and peer teaching with regard to basic arithmetic skills. After students have demonstrated that they can perform certain specified paper-and-pencil math calculations, they are notified that they must master related (and also specified) handheld-calculator skills before such and such a date—but that no class time will be used for that purpose. They are on their own, but if they wish, they can get help outside of class time at a computer clinic staffed entirely by (selected and trained) students. The teacher keeps notes on how students respond to the assignment and what the learning results are, and summarizes the episode for conference colleagues. (It would be necessary, of course, to keep an eye out for deleterious side effects that would compromise the success of some students. And it would be a good idea to be sure parents approve of the experiment.)

It is well to remember that a good alternative may not work terribly well the first time or two it is tried. Students as well as teachers need some time to get used to new procedures and expectations. Advice from experienced practitioners is highly desirable, for there are usually requisites that are not mentioned in the promotional literature. It would be helpful if enthusiasts who write articles about novel instructional formats disclose the difficulties as well as the triumphs (as, for example, advocates of cooperative groups might describe frankly the effort required to shape and sustain cooperation). Without this understanding that new approaches almost always take time to perfect, many good ideas may be discarded prematurely.

Traditional formats have their own side effects, inequities, and failures—but they are at least familiar and we more or less have learned to live with them.

As teachers in different grades try different formats with different content and in various situations, they may elect to adopt them as a more or less fixed part of their program, retain them to be used some years but not every year, or decline to make further use of them. As their efforts are documented and entered in the district's curriculum ideas file, other teachers, particularly those participating in the relevant network conference, will be encouraged to duplicate them, with the result that there may be a gradual increase in the use of seminars, projects, independent study, and peer teaching in the district. One sure result will be to increase the professional capability of the faculty.

**Alternative Time Patterns**

Much the same argument can be made for setting up computer-network conferences of teachers and administrators to explore various scheduling arrangements, with the objective of finding ways to pursue specific learning goals more successfully. Again, the purpose is to have some faculty members at every grade level and in every subject matter become knowledgeable about the advantages, limitations, costs, and risks associated with variable time scheduling that could accommodate (or be required within) highly valued curriculum blocks. Conference participants can build a new database of ideas for flexible time scheduling, and then teachers can volunteer to try out some of the ideas and document the experience.

Trying out alternative time arrangements may be more difficult to carry off than trying out alternative format ideas, but not prohibitively so, especially if several teachers take it on together. Some informal experimentation ought to be possible in most school situations. The basic rationale should be that a desirable unit or course requires a different time frame to realize its full effect. If the rationale seems cogent and results appear to be promising, then approval should be sought for trial on a larger scale, engaging administrators, parents, and students. Having to get approval gives another reason, in both cases, for spelling out in writing the nature and limits of the trial, the rationale for conducting the trial, and what information will be collected on its operation and effects.

Some of the conveniences of standard schedules are mentioned in the Chapter 2 section on schedule variations.

Some kinds of alternative time arrangements are relatively easy to test. In the elementary grades, self-contained classrooms are the norm. In principle, a teacher can configure daily instruction time in many different ways as long as the amount devoted to each subject over a longer time span—say, a semester—meets district and state requirements. “Block scheduling” on alternative days or semesters is becoming increasingly popular, and “modular scheduling” that allows for a variety of possible period lengths crops up every now and then.

Suppose, for example, that an average of 20 minutes per day of science is required. A teacher may consider apportioning that as two consecutive periods once a week, as 10 minutes a day plus a period once a week, or as nearly all day once every other week. The 10-minute-a-day alternative would best suit an

| Period | Monday | Tuesday | Wednesday | Thursday | Friday |
|---|---|---|---|---|---|
| 1 | Geometry | Geometry | Geometry | Geometry | Geometry |
| 2 | English | English | English | English | English |
| 3 | Spanish | Spanish | Spanish | Spanish | Spanish |
| 4 | PE | Semester Project | PE | Semester Project | PE |
| Lunch | Lunch | Lunch | Lunch | Lunch | Lunch |
| 5 | Music | Music | Music | Music | Music |
| 6 | Global Studies | Biology | Global Studies | Biology | Global Studies |
| 7 | Biology | Global Studies | Biology | Global Studies | Biology |
| 8 | Biology | Global Studies | Biology | Global Studies | Biology |

extended series of daily measurements; the all-day alternative would best suit an experimental problem that requires a long setup time or iterative trials. Such arrangements permit students to engage in both extended hands-on investigations and discussions of their findings, and they make it possible for elementary teachers who specialize in science to be used effectively. The same can be said for mathematics, art, and some other fields.

In middle and high schools, instruction is usually departmentalized, but that does not altogether preclude gaining experience with various nontraditional time arrangements. Following are two quite different possibilities, the first a fairly common experiment, the second very rare.

A variety of ways of partitioning time are described in Chapter 3.

In the first instance, arrangements are made for paired classes in which the same students are enrolled in two courses that meet in successive periods. (Science and mathematics have been thought to make a good combination, as do, for example, mathematics and physical education, science and technology, science and history, mathematics and social studies.) The two teachers then have ten periods a week to use in various ways, such as giving each student one double period and three single periods a week in each subject. This setup offers the possibility of extended instruction and of intersubject collaboration.

In the second instance, arrangements are made to commit one period each week to a genuine seminar (not “class discussion”) on a trade book. This takes place over a whole quarter or semester of a conventional one-period-a-day science course. To keep the seminar a reasonable size, the class is divided into three groups that explore different trade books on the same topic. The assumption is that such a seminar could enhance achievement of benchmarks in certain chapters of *Benchmarks for Science Literacy* (The Nature of Science, The Nature of Mathematics, The Nature of Technology, Historical Perspectives, Habits of Mind, and Common Themes) that deal with ideas that may not easily be learned through more closely managed instruction. But if the seminar format is initially more time consuming, some marginal topics could be eliminated from the course as a whole to accommodate the reduced amount of time available for conventional instruction.

In all such trials, keeping notes on the process and sharing results with colleagues should be part of the plan. Because the climate should be experimental, with the expectation that not everything will go smoothly the first time, there is no need for accounts to be as sunny and unblemished as in most journal articles. Difficulties and shortcomings can and should be reported, along with successes.

The *Designs on Disk* database of instructional formats can be of help in preparing reports and sharing them.

## Improving Assessment

It is essential to know something about how well each student is doing individually. Good teachers assess student performance frequently because effective teaching depends on having accurate and detailed information on what students do or do not understand. On a more summary level, parents need to know how well their children are doing in order to help fulfill their responsibilities properly; and eventually potential employers and admissions officers will base decisions about each student on his or her performance record. It is also important to estimate how effective the educational program is collectively. Schools, school districts, states, and the federal government periodically need to determine how well their populations—and defined subpopulations—of students are performing so they can make informed policy decisions about instructional practices and curriculum choices.

These three needs—to monitor the day-to-day learning of individual students, to summarize their individual progress and achievement, and to monitor the effectiveness of an educational program for the progress of defined populations of students—are often confounded. Judgments about individuals are ideally based on many sources of information, including what can be gleaned from quizzes and tests, homework assignments, essays and projects, portfolios of student work, teacher observation, and interviews. To report to third parties on the progress of students doesn't require nearly the amount of detail that the teachers themselves need. A single letter grade is usually too simple, greatly underrepresenting the very different possible profiles of knowledge and skills underlying it. But a summary profile of achievement, perhaps supplemented by brief commentary, may be satisfactory for many purposes (and can be followed up in more detail if desired).

To ensure fairness in judging individual students, all students should probably be subjected to the same assessment tools, or at least sets of equally difficult and representative tasks. But typically, individual tests cannot tap anywhere near all of what one would hope students have learned, but must settle for being a good sampling of that domain. And typically, the particular sample to be tested is unknown to students in advance, to prevent their study being narrowed to just that part of the domain.

To find out what a population of students has learned requires a very different approach. The simple, popular notion of giving the same test to every student in the population is monumentally wasteful. Acquiring complete and identical information on every student is far too expensive in terms of use of students' time and collection costs. However, reliable estimates about large groups can be made by sampling relatively small numbers of students—which, paradoxically, makes possible much better information. Rather than give the same test to 10,000 students, ten different tests covering ten different areas can be given to 1,000 students each. The sample size of 1,000 is enough to ensure a reliable estimate of what all 10,000 know on each test, and a ten times greater area of learning can be probed (or the same areas in ten times more detail) for the same investment of time. Moreover, a greater variety of performance can be assessed. In the same amount of time (say a one-hour examination), some students could get dozens of short-answer questions, others a dozen problems, and others a single higher-order problem-solving task.

Of course, results on such a large scale are not useful for reporting on individual students, because only a very small part of any one student's possible performance is tapped, and the different tests are not likely to be (nor need to be) precisely equivalent in difficulty. And, although such large-scale testing can produce far more information about the performance of a population of students, there is a new source of uncertainty: because results cannot be used to judge individual students, the motivation of students to show how well they can do would seem likely to be less. Because certain categories of students may experience a greater drop in motivation than others, not only the absolute but even the relative achievements may be misestimated. (Assessment experts are still working on the motivation question.)

## Alignment of Assessment and Curriculum

There are few topics in education that create more controversy than assessment does. In addition to disagreements over what techniques and instruments to use, there are differences over what the alignment should be among assessment, curriculum, and learning goals. The ideal, of course, is that curriculum and assessment be aligned both with each other and with specific learning goals. In a context of rapid change, however, there is inevitably some jostling about where change should begin. Commonly, the discussion revolves around two views—what may be called the "fairness view" versus the "leverage view."

The fairness view requires that assessment of students be aligned with the current curriculum. Students, teachers, and parents understandably want assessment to be fair, and fairness demands that students not be judged on what they have had no opportunity to learn. In this view, therefore, assessment should be aligned with the current curriculum—whatever it is. If and when the curriculum is changed, the assessment should change correspondingly. Thus, curriculum reform should lead and assessment should follow.

On the other hand, the leverage view requires that the curriculum be aligned with a new assessment. Many educators and citizens see assessment as a strategic lever for improving the curriculum. The proposition is that teachers will make (and parents and policymakers will welcome) whatever changes in the curriculum are necessary to help students do well on the tests. Change can more easily be made in assessment than in curriculum, and curriculum can be expected to adjust to assessment naturally. Thus, assessment reform should lead and curriculum reform will follow.

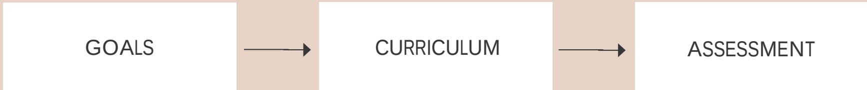
Both of those positions, obviously, imply a prior alignment. If assessment is to be aligned with the curriculum, what is curriculum to be based on? Or if assessment is to take the lead, from where does it obtain its authority? Curriculum and assessment are both means to ends, not ends in themselves. Hence, there is a third point of view that argues that curriculum and assessment both ought to be based on the same set of established learning goals. In this view,  curriculum and assessment are independently based on specific learning goals, and neither need be explicitly adjusted to the other. Schematically, these three points of view are parts A, B, and C in the box that follows.

Realistically, the completely independent approach is not likely to be satisfactory. One possible reason is that neither the curriculum nor the assessment will represent all the learning goals equally. If the balance arrived at in the curriculum is different from the balance arrived at in the assessment, there could be a significant mismatch.

## CURRICULUM AND ASSESSMENT RELATIONSHIPS

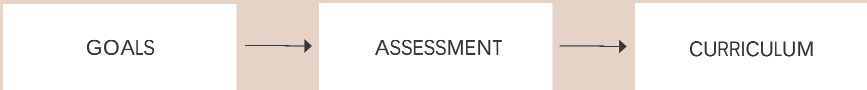
**A. Assessment follows curriculum**

An alternative to the assessment-first strategy of curriculum change is to begin with curriculum—which presumably has already been aligned with goals—and then align the assessment to what is actually taught. A danger here, however, is that the assessment will be tailored to fit incidental aspects of the curriculum that go outside of what was specified by the goals.



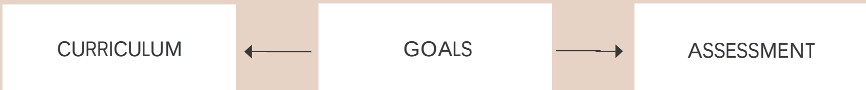

**B. Assessment leads curriculum**

Concern about the difficulty of changing curriculum leads to calling for changes in assessment as a means of inducing teachers to change what they teach. This instrumental use of assessment may arise from a passion about assessment methods per se, or from a concern about propagating a set of goals—to which the assessment itself would have to be already aligned.



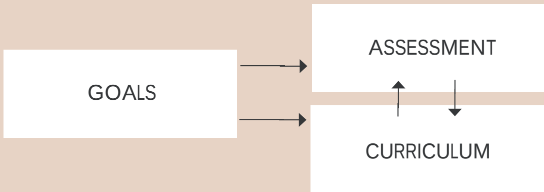

**C. Curriculum and assessment both derived from goals, independently**

Another view is that curriculum and assessment should both be aligned to the same specific learning goals, but by different groups. Differences in perceptions or preferences of the groups may produce uncomfortable discrepancies. Misalignments could also occur if the two groups focused on different subsets of the common goals.




**D. Curriculum and assessment both derived from goals, in parallel**

Congruence between curriculum and assessment can be improved by settling on specific goal emphasis for both in advance. Or, at least, drafts of curriculum and assessment could be reconciled to each other.



A second reason for not deriving curriculum and assessment independently is educators' uneasiness about making assessment of learning goals completely free of the particular context in which the goals were learned. Consider, for example, a benchmark on measurement skills in which the intention is that students could think about errors of measurement generally, in any context in which they are making measurements. But if measurement skills are learned largely in the context of studying oxygen levels in a local lake, teachers of that curriculum would be likely to assess student skill in those very measurements rather than, say, in a novel context of weather measurement. (Research does show, it is true, that learning does not easily transfer to new contexts.)

So a more practical plan would be to attempt to derive both the curriculum and the assessment from learning goals, but with each keeping an eye on the other and making accommodations from which each is likely to benefit. Regarding measurement skills, for example, the curriculum designers could build in more variety of measurement contexts to promote better transfer; and the assessment designers could provide students with a chance to show that at least they can measure in familiar contexts. This parallel alignment approach is illustrated by part D in the box opposite.

If a school district were starting out with a clean slate, the order of alignment would matter little in principle as long as ends (learning goals in this context) came first. In reality, though, few districts start with a clean slate: they already have a curriculum, though they may be in the process of changing it. On the other hand, few districts have a systematic process for assessing the effects of their curriculum. Thus the suggestions for action that follow concentrate on beginning to develop a school or school-district capability for curriculum evaluation that is designed to give good estimates of how well students in the district are reaching specified learning goals.

### Experimenting with Assessment

One can imagine a time when each course, large unit, or other block in a curriculum contains built-in assessment materials and procedures that have been carefully validated so that the block produces the student learning that the assessment measures. A district could then ascertain whether students are reaching the specified learning goals block by block (or at least at the grade-range checkpoints), and the sum of its findings would indicate what the curriculum as a whole has achieved. But good curriculum blocks appear to be more rare and more demanding to construct than may have been thought, so an adequate pool of blocks may be a long time in coming. In any

**Examples of assessment alternatives for a single benchmark on sampling**

*Multiple choice (30 seconds):*

The most important feature of a scientific comparison of two groups is

a. how group members were selected
b. the size of the groups
c. what percentage of the whole population do the groups represent
d. what kind of average is used to represent the groups

*Constructed response (5 minutes):*

Kim and Keisha had an argument over whether an MTV video was more popular among boys or among girls. Describe how they could choose groups of boys and girls to ask to find out who was right.

*Problem solving (5 days):*

Find out whether an MTV video is more popular among boys or among girls in your school.

case, it is possible to monitor a curriculum without having to monitor every course separately and completely.

It is important to distinguish the purposes and strategies of monitoring students as individuals from those of monitoring the effectiveness of the curriculum. Although the pool of tasks used for assessment may be the same, the tasks may be selected and employed in significantly different ways. For example, a much larger set of tasks can be used for curriculum monitoring by giving different students different subsets of tasks. The effectiveness of the curriculum can be estimated satisfactorily from the partial results and, since students need not be graded on the tasks, the different subsets need not be of precisely equal difficulty. The scoring need not be as elaborate for curriculum monitoring, since the number of students passing some threshold of success on each task may be all the information needed. For student monitoring, on the other hand, tasks can be selected more narrowly to fit the students' recent learning experiences and ability, yet interpreted more elaborately to diagnose the individuals' levels of understanding.

## Monitoring Classroom Learning

Curricula of the future are likely to demand that teachers are able to use a variety of conventional and nonconventional techniques for probing student learning. Skill in using multiple-choice tests, essays, portfolio analysis, oral interviews, and other approaches to assessment depends on understanding their advantages and limits, knowing when to use them, and being able to interpret their results correctly.

Although a high-quality assessment program requires long and reiterated development, a good way for a school district to develop its teachers' capability to employ such programs eventually is to have groups of teachers actually create samples of each kind of assessment and use them with students. The basic sampling process consists of the following five steps:

First, each group (subject and grade related), should select a single, specific learning goal of interest to the group and analyze it carefully to reach agreement on what it really means—and does not mean. If the learning goal is too general, the process may go off on different tangents. (Techniques for analysis are described earlier in this chapter.) The strand maps in *Atlas of Science Literacy* can be consulted in identifying precursors and connections to the key idea.

Second, subgroups of two or three teachers each should develop a different way of determining whether students have acquired the knowledge or skill defined by the selected benchmark (say, a set of objective short-answer items, an essay question, and a student interview). Research articles on student misconceptions may provide a starting point for identifying appropriate questions.

Third, at a suitable time, the subgroup teachers should apply their method to some students and use the responses to estimate the degree to which the learning goal has been reached by individual students. Each subgroup should write up the episode as a brief case study.

Fourth, the teachers should discuss their experiences and findings. The discussion should include how well the judges agree on the scoring and how well the assessment gets to the intended understanding or skill. Since discussion of the scoring can often lead to argument about just what each goal really means, the activity should increase teachers' understanding of the goals and their sensitivity to what constitutes student understanding.

Fifth, after discussion of the case studies, the group should select another benchmark—perhaps from those that strand maps show are closely related to the first one, or perhaps from a different content area entirely. For the new benchmark, each subgroup could change to a different assessment approach. Two or three cycles should be ample for all participants to become familiar enough with assessment techniques to make informed use of them immediately, if they wish, and to be prepared to employ them in redesigned curricula in the future.

Although some teachers may be inspired to develop more assessment tasks, others, who recognize how hard it is to invent good tasks, will settle for being better choosers and interpreters of tasks developed by specialists.

**Monitoring Curricula**

School districts that set up systems for monitoring the effectiveness of the curriculum usually conduct assessments of student learning at a few specified checkpoints—for example, at the end of grades 4, 8, and 11. For districts guided by *Benchmarks for Science Literacy*, the checkpoints of choice would be at the end of grades 2, 5, 8, and 12, and committees of teachers would be set up to develop some prototype assessments in science, mathematics, and technology for those grades. In the first year, a relatively few benchmarks would be selected and tests developed for them. At the appropriate time, a sample of students from all of the participating teachers would be tested, the data aggregated (so as not to be teacher specific) and then analyzed and discussed by the committee. In the second year, a larger selection of benchmarks would be made, and the process repeated. In this way, a district can gradually build a method for finding out how well its students are progressing toward science literacy, and, in the process, it can significantly increase the professional capabilities of its staff.

School districts usually have some staff who understand the design and statistics of large-scale assessments of learning and who can be called upon to provide leadership in building a capability for curriculum monitoring.

## Becoming Informed on Reform Movements

Finally, each school district should know what is going on nationally and statewide with regard to K-12 curriculum reform so that it can benefit from the concentration of resources and skills that are possible in large-scale projects. These projects issue reports that follow up on earlier reports and increasingly share their ideas and products more promptly and often on the Internet. Not every teacher needs to be so well informed. Nor is this to say that a district should feel compelled to introduce every new reform—or claimed reform—that surfaces. Rather, it should put itself in a position to decide knowledgeably which reforms to reject outright and which to consider further—maybe even which to adopt.

Committees should be set up in districts to track developments in a particular sector defined by subject matter, grade level, or aspect. Members would bring interesting ideas to the attention of colleagues (teachers, administrators, and school-board members), and serve as internal consultants on curriculum-reform matters. Nontheless, some of the curriculum committees should cut across grade levels and subject-matter fields, because many ideas for reform follow that pattern and because it helps to moderate the tendency of specialty groups to take a narrow view of the curriculum. Of course, the ability of these committees to be of real service depends on their having ready access to reports, newsletters, journals, books, and the Internet, along with office support and a budget for travel to conferences and to schools implementing reform of interest. In other words, building a professional capability for reform in a school district calls for support of work that includes but is not limited to classroom teaching.

The suggestions in this chapter, helpful as they may be, all imply more demand on teachers' time—a commodity already in painfully short supply. In the long run, changes in policy and financing are needed to make more teacher time available for planning and professional development. Yet, while there is no easy solution to the inadequacy of time for teachers' planning and professional development, such time as there is could be used more efficiently if it were better focused. In science, mathematics, and technology, that focus should be on promoting student learning of the carefully selected and coherent set of core ideas on which there is currently a wide national consensus. Project 2061's experience is that many dimensions of professional devel-

Project 2061's Professional Development Programs offer educators a variety of workshops and other services that can be customized to fit the long-term needs of each school or district.

opment can benefit from such a focus, including improved knowledge of subject matter and what aspects of it are most appropriate for students, better understanding of the general psychology of teaching and learning, skills in judging materials for instruction and assessment, and the value of collaboration with peers.

In the next chapter, consideration is given to how to reduce the number of topics taught. Abandoning less important topics will not directly make more time available for teachers, because the time that is freed up should be used to teach better the central ideas that remain. The tighter focus that will result from this effort will enable educators to make better use of the precious planning and professional-development time that is currently available.


*"Basically, we're all trying to say the same thing."*