JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 39, NO. 6, PP. 522–549 (2002)

How Well Do Middle School Science Programs Measure Up? Findings from Project 2061's Curriculum Review

Sofia Kesidou, Jo Ellen Roseman

Project 2061, American Association for the Advancement of Science, 1333 H Street, NW, Washington, DC 20005

Received 2 July 2001; Accepted 18 September 2001

Abstract: The purposes of this study were to examine how well middle school programs support the attainment of key scientific ideas specified in national science standards, and to identify typical strengths and weaknesses of these programs using research-based criteria. Nine widely used programs were examined by teams of teachers and specialists in research on teaching and learning. Reviewers found that whereas key ideas were generally present in the programs, they were typically buried between detailed or even unrelated ideas. Programs only rarely provided students with a sense of purpose for the units of study, took account of student beliefs that interfere with learning, engaged students with relevant phenomena to make abstract scientific ideas plausible, modeled the use of scientific knowledge so that students could apply what they learned in everyday situations, or scaffolded student efforts to make meaning of key phenomena and ideas presented in the programs. New middle school science programs that reflect findings from learning research are needed to support teachers better in helping students learn key ideas in science. The criteria and findings from this study on the inadequacies in existing programs could serve as guidelines in new curriculum development. © 2002 Wiley Periodicals, Inc. J Res Sci Teach 39: 522–549, 2002

Whereas curriculum materials (and in particular textbooks and their accompanying teacher's guides) are but one of the resources available to teachers, they have a major role in teaching and learning. Many teachers rely on them to provide some or all of their content and pedagogical content knowledge, and this is especially so when the teacher is a novice or is teaching outside his or her area of expertise (Ball & Feiman-Nemser, 1988; National Educational Goals Panel, 1994). Acknowledging their role in teaching and learning, both science education researchers and policy makers have called for systematic, research-based reviews of science curriculum materials as a means for improving their quality, influencing teacher practice, and supporting science education reform [Good, 1993; National Research Council (NRC), 1999, 2000b].

Correspondence to: Sofia Kesidou; E-mail: skesidou@aaas.org
DOI 10.1002/tea.10035
Published online in Wiley InterScience (www.interscience.wiley.com).

© 2002 Wiley Periodicals, Inc.

Reviews of science curriculum materials so far have had a narrow focus. They have typically considered isolated characteristics of curriculum materials, such as the range of topics covered in the material (e.g., Chiapetta, Sethna, & Fillman, 1993), the level of difficulty of the writing (e.g., Vachon & Haney, 1991), the use of illustrations (e.g., Holliday, 1990), attention to student diversity (e.g., Bazler & Simonis, 1991; Potter & Rosser, 1992), or the extent of inquiry focus (e.g., Eltinge & Roberts, 1993).

Only a few studies have explicitly looked at curriculum materials in the light of current theories of learning (Jimenez, 1994; Shiland, 1997). Findings from the learning sciences in the past 2 decades have not influenced systematically the evaluations of curriculum materials. This includes more general findings on knowledge organization in expertise, the role of prior knowledge (both conceptual and cultural) in learning, and conditions that facilitate the transfer of knowledge (NRC, 2000a). It also includes specific findings on characteristics of content area text that affect comprehension and learning as well as procedures for teaching students to learn from reading texts (Alexander & Kulikowich, 1994; Holliday, Yore, & Alvermann, 1994; Hynd, McWhorter, Phares, & Suttles, 1994), and content-specific instructional strategies and representations that support student learning (Chinn & Brewer, 1993; Guzzeti, Snyder, Glass, & Gamas, 1993; Scott, Asoko, & Driver, 1992; Smith, Blakeslee, & Anderson, 1993; White & Frederiksen, 1998). These findings have implications for the design of texts and teacher's guides, and can serve as the basis for criteria to judge the quality of curriculum materials (Good & Shymansky, 1986; NRC, 1999, 2000a, 2000b).

The study presented here is part of a long-term effort by Project 2061 to develop research-based criteria for reviewing science curriculum materials, and to carry out and make widely available reviews of curriculum materials based on these criteria. The study's purposes are to examine how well currently available middle school curriculum materials support the attainment of important scientific ideas specified in national science standards, and to identify typical strengths and weaknesses of these materials.

Conceptual Framework

Educators have different views on curriculum materials. They range from the view that curriculum materials need to prescribe to teachers exactly what to do, when to do it, and in what order (what was once called teacher-proof curriculum), to the view that curriculum materials are not needed at all, and the teacher must develop his or her own materials because he or she is the only one who knows the particular students' learning needs well enough to modify the classroom environment in response to these needs. In between, there are various other views, such as that curriculum materials are a necessary evil—some teachers (especially beginning teachers) are not yet adequately prepared to teach science—so curriculum materials may be needed until we have accomplished the job of large-scale teacher development; or that curriculum materials are useful reference materials—teachers need them as starting points from which they can put together materials of their own. We (like others) view curriculum materials as tools that allow teachers to do their best work with students. They provide a coherent science program for students based on the best thinking available in the field, and material that supports teachers in making more thoughtful and informed decisions about their own students science learning. Accordingly, our analysis of curriculum materials focuses on whether they meet the following essential requirements: their content primarily focuses on a coherent set of important, age-appropriate student learning goals; their instructional design effectively supports the attainment of the specified student learning goals; and the teacher's guides support teachers in helping students attain these goals.

Our efforts to develop a procedure to analyze how well curriculum materials meet these requirements are guided by a loose but unifying information-processing perspective of learning (Salomon & Perkins, 1998). Within this broad framework, the development of specific criteria has been pragmatic. Three successive groups of experts in science teaching and learning research, curriculum development, and kindergarten through Grade 12 (K–12) science teaching were asked to help formulate, negotiate, test, and refine a set of research-based criteria to judge the potential efficacy of curriculum materials (Roseman, Kesidou, & Stern, 1996). They were asked not to derive criteria from one approach but to look beyond labels, such as *constructivism*, and consider the implications of important findings from different approaches within the cognitive and science education learning research literature (such as those referred to in the introduction of this article). Particular attention was paid to content specific instructional strategies that have been shown to lead to effective science learning.

The criteria developed take into account both individual and social aspects of learning (Salomon & Perkins, 1998). For instance, a criterion holds curriculum materials accountable for providing opportunities for students to explain their ideas (and get feedback) both in writing and in small groups. Talk is important for sharing, clarifying, and distributing knowledge among peers. Analytical writing is an important tool for transforming rudimentary ideas into knowledge that is more coherent and structured. Furthermore, talk combined with writing appears to enhance the retention of science by students over time. Curriculum materials that attend to both are likely to enhance learning more than others that provide opportunities for either of these two language modalities alone (Pivard & Straw, 2000). The criteria are also consistent with equity concerns; they recognize individual differences and the need to provide explicit goals, hands-on experiences, and teacher modeling and scaffolding of science reasoning, curriculum attributes that serve diverse learners well (Lynch, 2000).

The Procedure in Brief

The Project 2061 curriculum-analysis procedure can be applied to a variety of K–12 science curriculum materials, ranging in length from a few weeks to several years. Briefly, the procedure requires reviewers to do the following.

Sample a Set of Learning Goals to Serve as the Basis for Analysis of the Curriculum Materials

A central proposition in the Project 2061 curriculum-analysis procedure is that curriculum materials are to be judged primarily in terms of how well they are likely to contribute to the attainment of specific learning goals. The procedure uses learning goals from national or state standards as the basis for evaluation. Because it would be impractical to undertake a thorough analysis of how well all learning goals specified in national standards are treated, sampling is needed. Our experience with a variety of curriculum materials has shown that analyzing a small but carefully chosen number of learning goals can provide a profile of the instructional strengths and weaknesses of a material as a whole. However, because curriculum materials sometimes differ in their treatment of particular content disciplines (physical, life, and earth science), learning goals from different disciplines need to be included in the sample.

Examine the Content of the Curriculum Material in Question to See How Well It Aligns with the Selected Learning Goals

This requires reviewers to look closely at the material to determine whether its content aligns with specific ideas in the learning goals that are sampled—not just whether the topics are

similar. Achieving alignment at the topic level is not difficult because most curriculum materials treat the same topics (heredity, weather, magnetism, etc.) that appear in national standards. However, materials may differ considerably in terms of which ideas within topics are treated. To be considered aligned with a learning goal, an activity or section of text should address the substance (rather than just the general topic) of the learning goal.

Determine How Well the Material's Instructional Strategies Support Student Learning of the Selected Learning Goals

The mere presence of specific content in a curriculum material does not ensure that students will learn that content. For learning to take place, curriculum material need to focus sound instructional strategies specifically on the ideas and skills that students are intended to learn. Reviewers rate each curriculum material on 22 criteria grouped under seven instructional categories. Criteria examine (among others) whether and how materials make their purposes explicit and meaningful to students, take account of student preexisting knowledge (both conceptual and cultural), provide a variety of phenomena and representations to make abstract ideas plausible and intelligible to students, guide student interpretation and reasoning, and help teachers create a learning-based classroom environment where all can succeed. Reviewers look for evidence that specifically defined characteristics, or indicators, for each criterion are met in the instructional guidance provided by a curriculum material in the student text or teacher's guide. A scoring scheme guides reviewers in using the indicators to assign ratings. Appendix A lists the specific criteria that serve as the basis for the instructional analysis. Appendix B lists the indicators and scoring scheme for the criterion Conveying Activity Purpose.

Summarize the Findings of the Content and Instructional Analyses

Having analyzed whether the content in the material matches the sampled learning goals and how well the instructional strategies in the material support students learning these goals, the final step in the procedure includes the development of criterion-by-criterion profiles that describe and rate how each sampled learning goal is treated in the material. Based on patterns in the profiles, conclusions are drawn about strengths and weaknesses of the material.

Method

The middle school science curriculum review focused on complete science programs (physical, life, and earth science materials from a publisher that span several grades) written specifically for the middle grades rather than on K–8 basal series or on supplementary or other materials that did not span Grades 6, 7, and 8. Nine complete science programs were selected. These include programs that are being widely used in school districts or states (identified by the Education Market Research Report, 1998), and through surveys distributed at meetings of the National Science Teachers Association), as well as three recent programs funded by the National Science Foundation (NSF). All but one of these programs were print-based. Appendix C lists the programs reviewed.

Three topic areas were selected to correspond to high priorities for physical, life, and earth science in the two major sets of nationally recommended learning goals (*Benchmarks for Science Literacy* (AAAS, 1993) and the *National Science Education Standards [NSES]* (NRC, 1996)) as well as in state frameworks: (a) the kinetic molecular theory (in physical science), (b) flow of matter and energy in ecosystems (in life science), and (c) processes that shape the earth (in earth science). Key ideas in these topic areas were identified from

relevant statements in both *Benchmarks* and *NSES*. These ideas represent key science concepts that any middle school program should cover. A list of these key ideas is available on-line at http://www.project2061.org.

There are two reasons why only discipline-based topics were selected rather than topics that deal with the nature of science, common themes (such as models or systems), or habits of mind. First, such topics are not equally well represented across state frameworks. Second, additional research and development efforts are needed to establish the reliability of the curriculum analysis procedure for these other important topics. The scarcity of information from research studies on how students learn these topics which reviewers can use to base their judgments makes such efforts particularly difficult.

The reviewers who participated in the analysis were evenly divided between experienced middle school classroom teachers and science education university faculty members who were knowledgeable about research on science learning and teaching. Reviewers were assigned to participate in the analysis of either the physical, life, or earth science topic area depending on their area of science content expertise. Two-member teams were formed, with one person on each team an experienced teacher and the other a science education researcher (whenever possible). Then, for each of the topic areas, the programs were examined by two 2-member teams with each team working independently of the other team. (Although all of the programs were examined by two teams, none of the teams examined all nine programs. Rather, depending on the time commitment they could make, each team examined two to five programs.)

Before the analysis, reviewers in each topic area received 7 days' training that had the following components: (a) clarification of the set of ideas that served as the basis of the analysis; (b) modeling the application of criteria, indicators, and scoring schemes using a range of examples from different curriculum materials; (c) practice using the Project 2061 curriculum analysis procedure with feedback on reviewers' understanding of the criteria, indicators, and scoring schemes both from peers and Project 2061 staff. A notebook with clarifications, scoring schemes, and a range of examples rated from poor to excellent on how well they addressed each criterion was available to reviewers during their training and during the subsequent analyses.

For each program, reviewers examined both the student text and teacher's guide to identify text segments, activities, teacher notes, assessments and other materials that addressed one or more ideas in each of the three topic areas. Next, they analyzed these chunks using the instructional criteria from the Project 2061 curriculum analysis procedure. In doing so, they analyzed the literal program—what instruction would be like if the teachers were to follow the suggestions in the teacher's guide literally. All teams used the same instructional criteria and scoring schemes to arrive at their ratings, and were instructed to produce a set of ratings on which both team members agreed. Teams were required to provide specific evidence from the materials to justify each of their ratings.

After analysis of all programs in each topic area was completed, the reports of the two teams that had worked independently on each program were compared, differences between the team ratings were noted, and the two teams were asked to reconcile their ratings and/or account for differences. Differences in ratings were due to several factors, which in turn determined how the differences were resolved. They most commonly reflected differences in the data used; owing to time constraints or because a specific feature of the program was not immediately recognized as relevant to a specific criterion, sometimes one or both teams did not identify all relevant places in the student text or teacher's guide that related to a specific criterion. In these cases, to reconcile the ratings, data from both teams were pooled, each item's pertinence to the criterion examined was reanalyzed, and the scoring guidelines were applied to the set of data which both teams agreed was pertinent. When, instead, the differences in ratings reflected

differences in interpretation of a criterion, indicator, or scoring scheme, a focused set of examples was provided to both teams to clarify further the intent of the criterion or indicator. Rarely, the two teams were not able to reconcile their differences despite further clarification, and in this case the final report included both teams' ratings (see Table 2). In the physical science topic area reconciliation occurred in 97%, in the life science topic area in 100%, and in the earth science topic area in 93% of the ratings.

For the purposes of this article, team ratings and reports were examined from two perspectives: whether the key ideas that served as the basis of the review were addressed in the programs; and what is the extent of instructional support the key ideas received for each of the criteria that served as the basis of the instructional analysis. This article identifies patterns across all three topic areas examined, illustrating them with findings mainly from the physical science topic area. Findings more specific to the life and earth science topic areas are presented elsewhere (Caldwell & Stern, 2000; Stern & Roseman, 2000).

Results

Are Key Ideas Addressed in the Programs?

In the context of the content analysis, reviewers examined whether programs addressed each key idea that served as the basis of the analysis within the topic areas kinetic molecular theory, flow of matter and energy in ecosystems, and processes that shape the earth. They further examined whether programs provided experiences with these ideas in one unit or distributed the experiences over several units (and grades), and if so, whether experiences were sequenced in a logical or strategic way and whether students were provided with opportunities to tie together the experiences they had with the same idea in different units.

On all three sets of ideas used to probe the conceptual development of important science ideas, 8 of the 9 programs reviewed did not differ greatly in their inclusion of the specific ideas. Each program addressed all (or nearly all) key ideas that served as the basis for the analysis. Only a few of the specific ideas were consistently absent, such as the idea that matter from food gets incorporated into the organisms' body in life science, or the idea that slow but continuous processes can, over a long time, cause significant changes on the earth's surface in earth science.

Although most key ideas were present in the programs, they were often buried between detailed, conceptually difficult, or even unrelated ideas, making it difficult for students to focus on the main ideas. This was true particularly for the life and earth science topic areas. For instance, programs present *photosynthesis* and *respiration* repeatedly (but briefly) in the context of detailed descriptions of cells, organisms, or ecosystems, but never develop them as instances of matter and energy transformation (a key idea in the standards documents). Programs also spend hundreds of pages on detailed descriptions of individual processes that change the earth (such as folding and faulting, volcanoes, earthquakes, wind erosion, water erosion, and glacial erosion) but rarely focus on the change that results from each process and how the different processes work together to change the surface of the earth (another key idea in the standards documents). Although sometimes the details included could have contributed to teaching the key ideas, more often they were not used to do so. This was true even in programs that listed the key ideas as main learning outcomes.

In 8 of the 9 programs examined, experiences related to the key ideas were distributed among several units (in 5 of these programs over two or three grades). However, in only 1 of these programs (and for only one topic area) were ideas encountered in progressively higher levels of sophistication. In all other programs, ideas were introduced without preparation and

were merely repeated rather than being explicitly revisited or clearly extended to new contexts, and the experiences that were provided for the same idea in different units were rarely connected to each other. For example, several programs include sections on molecular explanations of changes of state both in units on properties of matter (that appear earlier in the programs) and in units on heat and energy (that appear later) without linking the two presentations. Explanations of changes of state in the earlier units rely on the ideas that particles are in motion and increased temperature means increased molecular motion, although these ideas are explicitly introduced only in later units on heat and energy. Similarly, the concept of photosynthesis surfaces in multiple units (for example on simple organisms, animals, plants, oceans, and ecosystems), but each time it is mentioned as if the students have never encountered it before. Only one program alerted teachers to the individual units in which the same key idea was treated so that at least they can design their own ways of tying the different experiences together.

What Is the Extent of Instructional Support the Key Ideas Received?

Table 1 compares the mean ratings for the quality of instructional support for the nine programs reviewed in the three topic areas. In each topic area, the mean for each program was computed by averaging the reconciled criteria ratings over the 19 criteria in instructional Categories I–VI. In the few cases in which reviewers were not able to reconcile their ratings, the mean of the two individual teams' ratings was used as the criterion rating. None of the programs received a satisfactory mean rating (at least 2.0 points of a possible 3.0) in the physical, life, or earth science topic area. Rather, the mean ratings for the quality of the instructional support of the highest overall rated program (P9) were 1.42 [standard deviation (SD) = 0.67], 1.09 (SD = 0.69), and 1.11 (SD = 0.58) for the physical, life, and earth science topic areas, respectively. Table 2 presents the reconciled ratings by criterion for the physical science topic area.

Table 1
Mean instructional ratings of programs reviewed

	Physical Science (Topic Area: Kinetic Molecular Theory)								
	P1	P2	P3	P4	P5	P6	P7	P8	P9
Mean	1.08	0.76	1.41	0.53	1.11	1.13	0.66	1.21	1.42
Range	0–2	0–1.5	0.5–2.5	0–1.5	0–2.5	0–2	0–1.5	0.5–2	0–2.5
SD	0.58	0.48	0.68	0.49	0.70	0.60	0.44	0.45	0.67
	Life Science (Topic Area: Flow of Matter and Energy in Ecosystems)								
	P1	P2	P3	P4	P5	P6	P7	P8	P9
Mean	0.62	0.79	[a]	0.33	0.61	1.11	0.36	0.63	1.09
Range	0–1	0–1.5		0–1.5	0–1.5	0–2.5	0–1	0–1.5	0–2.5
SD	0.38	0.56		0.41	0.46	0.70	0.38	0.44	0.69
	Earth Science (Topic Area: Processes That Shape the Earth)								
	P1	P2	P3	P4	P5	P6	P7	P8	P9
Mean	0.70	1.0	0.76	0.55	0.45	1.18	0.52	0.72	1.11
Range	0–1.5	0–2	0–2	0–1.5	0–1.5	0–2.5	0–1	0–1.5	0–2
SD	0.41	0.58	0.67	0.50	0.50	0.76	0.34	0.45	0.58

Note. References to the programs reviewed (P1–P9) are listed in Appendix C.
0–1 = poor; 1.5 = fair; 2 = satisfactory; 2.5 = very good; 3 = excellent.
[a]P3 does not address the key ideas that served as the basis of the analysis in life science.

Table 2
Physical science: Instructional ratings of programs reviewed by instructional category or criterion

Instructional Category/Criterion	P1	P2	P3	P4	P5	P6	P7	P8	P9
I. Providing a sense of purpose									
Conveying unit purpose	1	1.5	1.5	1.5	1	2	1	1	1
Conveying lesson purpose	1.5	1.5	2	0.5	1	1	0.5	1	1.5
Justifying activity sequence	1	1	2–3[a]	0.5	1	2	1	1	1.5
II. Taking account of student ideas									
Attending to prerequisites	1	1	1	0	1.5	0.5	1	1	1
Alerting to commonly held student ideas	0	0.5	1	0	0	0	0.5	0.5	0.5
Assisting teacher in identifying student ideas	0	1	1	0	0.5	1.5	0	1	2
Addressing commonly held ideas	0	1	0.5	0	0	0	0	0.5	0.5
III. Engaging students with relevant phenomena									
Providing variety of phenomena	1.5	0.5	1	1	2	1	1	2	2
Providing vivid experiences	1.5	1	2	1	2.5	1.5	0.5	2	2
IV. Developing and using scientific ideas									
Introducing terms meaningfully	1.5	1.5	2.5–3	1	2.5	1.5	1.5	2	2
Representing ideas effectively	2	1	1	0.5	1	1	1	1	1
Demonstrating use of knowledge	1.5	0.5	1–3	1	1	0.5	0	1	0
Providing practice	1.5	0.5	1	1	1.5	1	1	1.5	2
V. Promoting student thinking									
Encouraging students to explain their ideas	1	1	2	0.5	1	2	0.5	1.5	2
Guiding student interpretation	1.5	0.5	2	1	1	1.5	0.5	1	1.5
Encouraging students to think about what they have learned	1.5	0	1–2	0	0.5	1	0	1	1

Note. References to the programs reviewed (P1–P9) are listed in Appendix C.
0–1 = poor; 1.5 = fair; 2 = satisfactory; 2.5 = very good; 3 = excellent.
[a]Independent ratings of the two reviewer teams for the four criteria for which no reconciliation occurred.

The following discussion presents in detail some of the typical strengths and weaknesses of the programs examined focusing on instructional Categories I–V (see Appendix A). Given space limitations, this article focuses on results from instructional Categories I–V that relate to the basic program (student text and teacher's guide) and are examined with regard to specific ideas. Category VI relates mostly to programs' assessment packages that are typically not included in the basic programs. Category VII examines, independently of specific learning goals, whether programs help teachers to create a supportive learning environment in which all can succeed. Detailed reports on each programs' performance in all instructional categories (including Assessing Progress and Enhancing the Science Learning Environment) are available at http://www.project2061.org. To facilitate interpretation of the findings, the discussion of each instructional category begins with a brief description of the category's intent.

Identifying and Maintaining a Sense of Purpose. The positive effects of providing students with a clear understanding of the purpose, goals, and content of an activity have been repeatedly confirmed by educational research (Boulanger, 1981; Wise & Okey, 1983). If students are to derive the intended learning benefits from engaging in an activity, their interest in or recognition of the value of the activity needs to be motivated (Blumenfeld, Soloway, Marx, Krajcik, Guzdial, & Palincsar, 1991; Malone & Lepper, 1987). Moreover, students need to understand the activity's goals and what these imply about how they should respond to the activity (Tasker & Freyberg, 1985). This category includes criteria to determine whether the curriculum material attempts to make its goals clear and meaningful to students, either by itself or by instructions to the teachers.

The units related to each topic area examined were analyzed to identify text segments, questions, or commentary in the teacher's guide that would convey to students (or ask teachers to convey) the overall purpose and direction of the unit. In addition, individual activities (including reading assignments, demonstrations, and labs) were analyzed to assess whether their purpose and their relationship to other activities in the unit were adequately conveyed. Table 2 presents the ratings that each program received for the physical science topic area. Table 3 presents the indicators used to assess the unit and activity purposes of programs and the results obtained for the physical science topic area.

All programs examined provided some purpose for the units of study (unit or chapter); however, only for one (NSF-funded) program were the purposes judged to be satisfactory (Table 2). In several programs, purpose statements used abstractions with which students would not be familiar at the beginning of the unit [e.g., "Point out that in this unit, they will be studying the structure of matter and how its structure determines its characteristics and behavior" (McLaughlin & Thompson, 1997, p. 211)], or related to experiences that are likely to be too commonplace to be interesting to middle school students. For example, after presenting a scenario about students' experiences on the beach (e.g., cooling off in the water after lying in the hot sun, buying ice cream and watching it melt), a text asks, "What causes all these changes? You'll find out in this chapter with the help of the kinetic molecular theory" (Aldridge et al., 1998, p. 208).

In six programs, unit openers included questions that framed only a small part of the unit, making it unlikely that interest (even if aroused) would be sustained over the duration of the unit. For example, the beach scenario in the previous example was used to frame a chapter on properties of solids, liquids, and gases; changes of state; pressure; gas laws; and absolute temperature. None of the programs asked students consistently to return to the stated purpose at the end of units, providing closure and thus sustaining interest and enhancing the perceived value of the units (Blumenfeld et al., 1991).

Table 3
Physical science: Number of programs that met indicators for criteria "unit purpose" and "activity purpose" (n = 9)

Unit purpose		Activity purpose	
Indicators	No.	Indicators	No.
1. Purpose presented	9	1. Purpose presented	6
2. Purpose comprehensible	3	2. Purpose comprehensible	5
3. Purpose interesting or motivating	1	3. Students think about purpose	0
4. Students think about purpose	0	4. Activity purpose related to unit purpose	0
5. Lessons consistent with purpose	3	5. Activity connected to what comes before or after	1
6. Material returns to purpose	0		

At the individual activity level, six programs consistently attempted to convey the purpose of activities to students (Table 3); in only one (NSF-funded) program were the purposes judged to be satisfactory (Table 2). This program presented purposes for text segments and explorations in ways that were likely to make sense and be meaningful to students who had not yet studied the topic and were not familiar with the scientific vocabulary. In addition, it made an attempt to connect individual activities to each other by describing to students what they had learned so far and what they needed to learn or do next. The following is an example:

> Now you know several parts of the particle model: All materials are composed of particles; particles in hot things are moving faster than particles in cold things; particles attract each other, and particles in different materials are uniquely arranged and have unique interactions. The next few investigations will give you the chance to use all of this information to create scientific explanations for more phenomena. (...) In this investigation, you will determine some of the properties of a very curious material. Then you will have a chance to develop your own scientific model by using your own ideas and what you already know of the particle theory. Your model will need to explain the properties of the strange material. You also will see what it is like for someone to ask you to test your model and then have to decide how to change your model to fit the results of the tests. (BSCS, 2000, pp. 218–219)

Several programs included brief objectives for text segments and explorations but the emphasis was on what students would do, but not why ("systematically investigate and describe matter that you cannot see"; Atwater et al., 1995, p. 38); used terms that students are not likely to know before they study the topic ("Make models illustrating the gas laws"; Science Insights, p. 140) ("Account for the energy of the heats of fusion and vaporization in state changes"; McLaughlin & Thompson, 1997, p. 224); or did not reflect adequately the actual purpose of the activity, e.g., the question "Do hot things move?" was used to frame a demonstration of candy dissolving faster in hot water than in cold water, which was intended to illustrate that increased temperature means increased molecular motion (Aldridge et al., 1998, p. 209). All programs presented the purpose of activities in statements or at best in questions used in a rhetorical way, making it unlikely that students' active, generative cognitive processes would be engaged. No program engaged students in thinking themselves about the purpose of the activities (why they should do them, or how they are connected to the overall purpose of the unit).

Taking Account of Student Ideas. Fostering understanding in students requires taking time to attend to the ideas they already have, both ideas that are incorrect and ideas that can serve as a foundation for subsequent learning (Eaton, Anderson, & Smith, 1984; McDermott, 1991; Minstrell, 1984; Roth, 1991; White & Frederiksen, 1998). Such attention requires that teachers are informed about prerequisite ideas and skills needed for understanding a new idea and what their students' initial ideas are (both conceptual and cultural), in particular the ideas that may interfere with learning the scientific ideas. Teachers can help address students' ideas if they know what is likely to work (Anderson & Smith, 1984; Roth, 1984). Curriculum materials that alert teachers to their students' likely misconceptions, suggest strategies for identifying and dealing with them, and incorporate appropriate strategies to take account of students' ideas greatly enhance teachers' effectiveness in promoting student understanding (Bishop & Anderson, 1990; Brown & Clement, 1992; Eaton et al., 1984; Lee, Eichinger, Anderson, Berkheimer, & Blakeslee, 1993; Songer & Linn, 1991).

Seven of the nine programs examined included claims that they attended to students' prior knowledge (for example, ''The role of past experiences and preconceptions as either an aid or an obstacle to learning is incorporated in the material's design'') and five programs included specific features (labeled as Revealing Preconceptions, Prior Knowledge and Misconceptions, or Addressing Misconceptions) intended to help teachers take account of student ideas. Despite this attempt to pay attention to students' prior knowledge, only one program received a satisfactory rating for one of this category's criteria for the physical science topic area (Table 2). Programs tended to focus on trivial misconceptions rather than student ideas that interfere significantly with student learning.

For instance, one criterion probes whether curriculum materials alert teachers to commonly held student ideas, both troublesome and helpful. Table 4 presents the indicators used to assess this criterion and the results obtained for the physical science topic area. Although in six programs places were recorded in which the teacher's guide alerted teachers to specific student misconceptions (Table 4), in four of these programs the misconceptions related to peripheral ideas rather than the central important ideas the units were trying to develop. For example, there is extensive literature on students' conceptions and epistemological commitments that interfere significantly with students' learning the particle model and the kinetic molecular theory. Students think that particles (atoms or molecules) are in substances (rather than that substances

Table 4
Physical science: Number of programs that met indicators for criteria "alerting to commonly held ideas" and "assisting teacher in identifying student ideas" ($n = 9$)

Alerting to Commonly Held Ideas	No.	Assisting Teacher in Identifying Student Ideas	No.
1. Presents commonly held ideas	6	1. Includes questions identified as eliciting student ideas	5
2. Commonly held ideas relate to key ideas	2	2. Questions relate to key ideas	2
3. Clarifies or explains commonly held ideas	0	3. Questions likely to be comprehensible to students	2
		4. Includes questions that ask for predictions or explanations (rather than focus on factual information or use of terms)	1
		5. Probes beneath initial responses	0
		6. Helps teachers interpret student responses	0

are made of molecules), and/or that there is something (e.g., air) between particles (Brook, Briggs, & Driver, 1984; Nussbaum, 1985); attribute macroscopic properties such as hardness, hotness, coldness, expansion, and physical state to particles (Lee et al., 1993); and favor explanations other than reductionistic and system-based ones of the atomic and kinetic molecular theories (Albanese & Vicentini, 1997; Hesse & Anderson, 1992). However, none of these findings are described for the teachers' benefit in any teacher's guide, nor are they used in any strategic way in student activities. Only one program alerts teachers that "You might discover that they [students] hold the common misconception that the individual particles are getting larger as they get hotter" (BSCS, 2000, p. 191), but it does not explain the misconception as one aspect of students' tendency to attribute macroscopic features to microscopic particles. In contrast, four programs alert teachers to trivial misconceptions, such as that many students erroneously call the cloud that emerges from a teakettle or hangs over a bathtub "steam."

A similar focus on detailed facts rather than on key ideas was documented in programs' attempts to help teachers identify their own students' ideas. For example, five programs included questions intended to elicit student ideas in the physical science topic area (Table 4). However, the questions were mostly factual and often peripheral to the main ideas the units were trying to develop, and/or were expressed in ways that are not likely to be meaningful to the students who have not yet studied the topic. For instance, before instruction of the atomic and kinetic molecular theories, one program suggested the following questions to assess student ideas: "What is the name given to the smallest pieces of matter? What does every object in the universe have in common? How do you know air is matter?" (Atwater et al., 1995, p. 6).

Furthermore, no program provided help for teachers to interpret student responses or specific guidance on how to make use of the information. Without explicit knowledge of the problematic ideas and reasoning that students have, it is unlikely that teachers will address student difficulties (Smith et al., 1993). The student texts did not explicitly address student ideas (Guzzetti et al., 1993), nor did teacher's guides suggest activities that could help students with their difficulties, such as challenging student ideas in appropriate ways (Chinn & Brewer, 1993); contrasting student ideas with the scientifically accepted ones (Smith et al., 1993); or extending commonly held ideas that have limited scope (Brown & Clement, 1992).

Engaging Students with Phenomena. Effective teaching provides students with opportunities to relate the scientific concepts they are studying to a range of appropriate phenomena through hands-on activities, demonstrations, audiovisual aids, and discussions of familiar phenomena (Anderson & Smith, 1987). Appropriate phenomena help students to view scientific concepts as plausible, or enhance students' sense of usefulness of scientific concepts (Anderson & Smith, 1987; Champagne, Gunstone, & Klopfer, 1985; Strike & Posner, 1985).

Units related to each topic area examined were analyzed to identify phenomena that provide direct support for each key idea (rather than being just peripheral to it). We then examined whether these phenomena were explicitly linked to the key idea, whether they were presented to students in a vicarious way, and whether there was a sufficient variety of phenomena to support the key idea. Although for the topic area "kinetic molecular theory" three programs (one of which was NSF-funded) received a satisfactory rating (Table 2), no program supported explicitly and systematically the key ideas related to the topic areas "flow of matter and energy in ecosystems" and "processes that shape the earth" with phenomena. This reflects the lack of focus of programs on the key ideas in the life and earth science topic areas noted earlier. Although all programs claimed to be activity-based and indeed included several hands-on activities in each chapter, activities only occasionally targeted key ideas and were often add-on

Table 5
Phenomena related to key idea "plants use energy from light to make energy-rich sugars"

Phenomena Directly Support Key Idea	Phenomena Are Not Adequately Linked to Key Idea	Phenomena Focus on Peripheral Ideas
Transformations of light energy to other forms of energy (e.g., chemical changes to photographic emulsions)		Plants move toward a light source
Formation of sugar or starch occurs in the presence of light but not in its absence	Students test for starch in leaves exposed to light and leaves covered. Although this could be discussed as evidence in support of the key idea, the only questions posed require noting the results and responding to why *photosynthesis* is a good name for the process plants use to make food. The answer in the teacher's guide, "It means making something of light," does not express the key idea	The rate of photosynthesis increases as the light intensity increases
The colors of light absorbed by chlorophyll are the same colors that result in plant growth	After noting that light passing through chlorophyll extract does not contain some colors of light (a phenomenon that could be discussed as evidence in support of the key idea), the only question students are asked is, "Why do you think leaves containing chlorophyll appear green?"	Green leaves are composed of several pigments

features that were neither well integrated with the rest of the text nor explicitly linked to the key ideas.

For example, the first column in Table 5 lists phenomena that directly support the key idea that plants use energy from light to make energy-rich sugars. Such phenomena were rarely included in the programs examined. The second column lists examples from programs that included relevant phenomena but missed the opportunity to link them adequately to the key idea. The third column lists phenomena that are peripheral to the key idea but were typically included in the programs examined.

Developing and Using Scientific Ideas. Providing students with experiences with phenomena (see previous instructional category) is not sufficient for them to develop an understanding of the principles and concepts of science. Scientific ideas cannot be inferred directly from phenomena; they need to be introduced explicitly and taught to students (Driver, 1983; Driver, Asoko, Leach, Mortimer, & Scott, 1994). Multiple representations are needed to make the ideas intelligible (Champagne et al., 1985; Spiro, Feltovich, Coulson, & Anderson, 1989; Strike & Posner, 1985). Students need help to see how the ideas can be used in activities that science-literate individuals are expected to perform, such as describing and explaining phenomena, solving practical problems, and considering alternative positions on issues (Anderson & Roth, 1989). Students initially learn new ideas in specific task contexts and often have difficulty transferring what they have learned in other contexts. Therefore, students need opportunities to apply ideas in a variety of contexts. Real competence comes only with extensive practice with feedback (Anderson & Roth, 1989; Ericsson, Krampe, & Tesche-Romer, 1993; National

Research Council, 1994). This category includes criteria that assess whether materials represent scientific ideas in a variety of ways that are accessible and intelligible to students, model the use of scientific ideas in describing and explaining phenomena, and provide tasks for students to practice using scientific ideas in a variety of situations.

We analyzed the units that related to each topic area to identify representations (drawings, diagrams, graphs, analogies and metaphors, models, and simulations) pertinent to each key idea. We then examined whether the representations were accurate and comprehensible, and whether they were linked to the idea or process that they represented. In the physical science topic area, only one program received a satisfactory rating (Table 2). A consistent problem with all but two of the programs examined was the inaccuracy and misleading nature of the representations included. For example, in teaching the kinetic molecular theory, six of the nine programs included illustrations showing particles of liquids and gases in containers against colored backgrounds (as if they were floating in a substance). Students who do not distinguish between model and observation, conceive matter as continuous, and have difficulty imagining empty space may take this illustration literally and think that particles are contained in substances (rather than that substances are made of particles), a common student misconception (Andersson, 1990). Of course, because representations can highlight only some aspects of an idea or process, no representation can be completely accurate. Nevertheless, curriculum materials could address this issue by asking students to analyze how the representations included are like and unlike the thing being represented (Treagust & Harrison, 2000). None of the programs examined asked students to look critically at representations.

Before students can be expected to use scientific ideas, they need to see what is expected. For example, it may be useful to use the idea that molecules are constantly moving to explain what happens when we open a bottle of perfume, before we expect students to explain what happens when we cut into a lemon or smell trash. For each key idea, places were identified in which the program explicitly modeled its use, and we then examined whether the performance was step-by-step and whether criteria for judging a good performance were identified. None of the programs examined showed students explicitly how to use scientific ideas. The materials funded by the NSF rarely included explanations of phenomena in the student text. When explanations were included, they were rarely step-by-step. For example, explanations of changes of state included in programs did not describe explicitly the arrangement, motion, and interaction of particles in the different states of matter, and how these change as the temperature increases or decreases. Instead, the following explanation for the solidification of candle wax was typical: ''This candle has just been blown out. The liquid wax at the top of the candle is cooling. As its temperature drops, its molecules move more slowly. Eventually, they clump together and the liquid changes to a solid.'' In all programs, teacher's guides included responses to questions asking for explanations in the student text. However, the responses given are too brief to be helpful to teachers who, despite years of experience, may have difficulty modeling scientific answers to questions that ask for explanations because they do not have adequate models themselves (Heller, 2001; Hollon, Roth, & Anderson, 1991). Moreover, because teacher's guides do not instruct teachers to use the answers to model the use of knowledge, the teachers might use them only to correct student papers, and not let the students hear or read the correct answers.

Given that students learn well what they practice doing, an important part of learning science consists of having several opportunities to exercise one's understanding of science. For each key idea, we examined whether programs provided opportunities for students to practice its use in a variety of contexts, whether programs provided a sequence of tasks for practice in which the complexity of the tasks was progressively increased, and whether programs

presented students first with opportunities for guided practice with feedback and then with practice in which the amount of support was gradually decreased. Having students practice using ideas was not a focus of programs (both traditional textbooks and NSF-funded programs). Unlike their emphasis on the importance of hands-on activities, the introductions to teacher's guides rarely emphasized the importance of giving students the opportunity to practice using the scientific ideas developed. Whereas traditional programs typically include fill-in-the blank types of questions or questions for which students can copy the answer from the text at the end of each section [e.g., "Contrast the movement of particles in solids, liquids, and gases" (DiSpezio, Linner-Luebe, Lisowski, Skoog, & Sparks, 1997, p. 154), "How does heat energy cause matter to change phases?" (DiSpezio, Linner-Luebe, Lisowski, Skoog, & Sparks, 1997, p. 226)], the NSF-funded programs included one or two questions that asked students to apply their understanding in everyday contexts. For example: "Imagine you and your friends are a group of water particles in a block of ice that is floating south along the east coast of America. Use ideas about particles in solids, liquids, and gases you have learned in this chapter to help you write words for a song. The song should tell the story of what happens to you over the period of a year" (The Prime Science Education Group, 1998, p. 65). However, the small number and variety of questions included, and the fact that no provisions were made for how students would receive feedback on their responses, make it unlikely that most students would achieve mastery without significant additional teacher support.

Promoting Student Thinking about Phenomena, Experiences, and Knowledge. Engaging students in experiences with phenomena (Instructional Category III) and presenting them with scientific ideas (Instructional Category IV) does not lead to effective learning unless they are given time, opportunities, and guidance to make sense of the phenomena and ideas (Driver, 1983; Driver et al., 1994). No matter how clearly materials may present ideas, students will make their own meaning out of it (Anderson & Smith, 1984; McNamara, Kintsch, Songer, & Kintsch, 1996). Carefully chosen and sequenced questions and tasks are necessary to scaffold students' attempts to construct the intended meaning of experiences or presentations of ideas (Anderson & Roth, 1989; Anderson & Smith, 1987; Arons, 1990). When students make their thinking about experiences and ideas overt to themselves, the teacher, and other students, their thinking can be examined, questioned, and shaped (Chi, 2000; Linn & Burbules, 1993; NRC, 2000a; Roschelle, 1992). Feedback on one's thinking is important for learning, especially when it focuses on understanding rather than on memory of procedures or facts (NRC, 2000a).

Table 6 presents the indicators used to assess two of the criteria in this category: whether materials provide students with opportunities to explain their ideas and reasoning, and whether materials include scaffolded questions and tasks that guide students' interpretation and reasoning about phenomena and ideas. Four programs (two NSF-funded) included routine opportunities for students to express their ideas with regard to key ideas and phenomena, both in classroom discussions and through use of journal writing and small group work (which provide opportunities for *each* student to express his or her ideas). Although most other programs included features such as Discussions Strategies, Portfolio, Embedded Assessment, or Science Journal, questions in these features merely required restating phrases in the text rather than encouraged students to express their own thinking, or they targeted mostly peripheral issues rather than key ideas. An example follows: "Journal Activity: You know that water is a liquid. You also know that when it is frozen, water can also be a solid. Make a list of all the kinds of things that you can do with liquid water. Make a second list of things you can do with solid water. Enter your lists in your journal" (Maton, Hopkins, Johnson, LaHart, Warner, & Wright, 1997, p. 61). None of the

Table 6
Physical science: Number of programs that met indicators for criteria "encouraging students to explain their ideas" and "guiding student interpretation and reasoning" (n = 9)

Encouraging Students to Explain Their Ideas	No.	Guiding Student Interpretation and Reasoning	No.
1. Encourages students to express their ideas	4	1. Includes relevant questions	6
2. Encourages students to clarify, justify, and represent their ideas	0	2. Questions frame important issues, help relate students' ideas to phenomena presented and scientific ideas, or help relate phenomena to scientific ideas	1
3. Opportunities for each student to express ideas	4	3. Includes scaffolded sequences of questions	0
4. Includes suggestions on how to provide explicit feedback	0		
5. Feedback signals state of student understanding (rather than progress in memorizing facts)	0		

programs provided routine opportunities for students to clarify, justify, and represent their ideas, which make the activity of expressing ideas more effective (e.g., Woloshyn, Willoughby, Wood, & Pressley, 1990). None of the programs made suggestions for when and how students would get feedback from peers, the teacher, or other resources. Rather, they either merely provided the correct answer to questions or stated that student answers to questions "will vary." None of the programs consistently provided help for teachers to anticipate student responses and conceptions on which they are based, or suggested how teachers could respond to them, or how student ideas may be further developed. Such information could help prepare teachers for some of the uncertainties of teaching, although, of course, it cannot provide definite prediction of specific student responses (Ball & Cohen, 1996).

Specific suggestions to guide students' interpretation and reasoning about the activities in which they engage were also inadequate in most programs (see, for example, Table 6). Sometimes a single general question was posed at the end of an activity in the student book about what the activity showed or what conclusions could be drawn from it. In the physical science topic area, only one program included consistently focused questions to encourage students to relate what they were observing in hands-on activities to key ideas. No program included questions that anticipated common student misconceptions, focused on contrasts between student misconceptions and scientific alternatives, or encouraged students to relate their own ideas to phenomena observed or ideas presented. Such questions are likely to help students identify and resolve conflicts and misunderstandings in their representations (Chi, 2000), have been shown to be effective in promoting student learning (Smith et al., 1993), and can be derived from the research on learning about particular topics (Roth, 2001). No program included carefully structured sequences of questions to scaffold students' drawing inferences from observations. Instead, in most programs students were required to make large leaps to get from evidence to conclusions, which may not only confuse but also discourage them from further study of science (Arons, 1990). A typical example follows: Students are asked to put a teabag into water and observe what happens (The Prime Science Education Group, 1998, p. 56). Then they tear open a teabag and with a microscope examine the paper it is made from. A single question is provided—"What does this suggest to you about the pieces of tea that spread out in

the water?''—to guide student interpretation of the activity, which is too general to help students connect what they observe to the idea the activity is intended to develop (''the particles of dissolved solids are too small to be seen'').

Summary and Discussion

The findings presented here indicate that none of the middle school programs examined are likely to contribute to the attainment of the key ideas that served as the basis of this study's analysis. First, the programs' content does not focus on the key ideas. Rather, details irrelevant to learning key science ideas are interspersed with details distracting from key ideas; details that are peripheral but could be used to support key ideas are not linked to the underlying key ideas; and experiences provided for the same key idea are distributed over several units, but not linked with each other.

Second, the programs' instructional design does not support the attainment of the key ideas. Programs are particularly deficient in providing coherent explanations of real-world phenomena using key science ideas, and building on students' existing ideas or helping them overcome their misconceptions or missing prerequisite knowledge. No program received a satisfactory rating for these criteria (Table 2). Programs claim (and attempt) to include research-based innovative practices, such as Embedded Assessment or Science Journals. However, they do not consistently focus them on key ideas but rather on peripheral ideas or other details. Components that claim to attend to student prior knowledge and misconceptions are typically add-ons that refer mostly to students' lack of knowledge of specific factual information rather than efforts to identify, clarify, and suggest how to deal with deep-rooted beliefs likely to interfere significantly with learning.

Third, the programs' support to teachers in helping students attain the key ideas is minimal. Teacher's guides provide minimal help for teachers to scaffold students' attempts to make sense of key ideas and phenomena, and no help in interpreting students' responses to the programs questions and activities and in providing appropriate feedback.

All programs examined display these shortcomings, although individual programs have limited strengths. Science Plus, the program that received the highest overall rating, received a satisfactory rating for 5 criteria (of the 16 included in Instructional Categories I–V) in the physical science topic area. Still, teachers who use this program will need to draw on a variety of resources to compensate for its weak areas. For example, although *Science Plus* was rated satisfactory for providing a variety of phenomena to make key ideas plausible to students, the program's support for helping students to make sense of these phenomena was rated fair (Guiding Student Interpretation and Reasoning). Although it was rated satisfactory for providing opportunities for students to practice using key ideas to explain everyday situations, it was rated low in modeling how to use key ideas to construct explanations of everyday phenomena (Demonstrating Use of Knowledge).

National Science Foundation–funded programs (P3, P5, and P6 in Table 2) display limited strengths compared with other programs. They usually reduce the amount of content included (especially content that is developmentally inappropriate for middle school students); limit the number of technical terms (potentially contributing to the readability of their texts) and link the terms used to relevant experience; and provide opportunities for students to explain their own ideas. However, like the non–NSF-funded programs, NSF-funded programs do not take account of students' prior knowledge and beliefs, lack coherent explanations of real-world phenomena using key science ideas, lack opportunities for students to practice developing such explanations, and provide limited help for teachers to scaffold students' attempts to make sense of phenomena and ideas.

Each program examined was generally weak in its support for all three of the topic areas that served as the basis of this analysis (Table 1). This is not surprising given that programs include identical instructional strategies for all the content areas they address. Small differences in ratings of the three topic areas examined appear to be due to differences in how well materials focused on the specific key ideas that served as the basis of the analysis. Across the board, the programs' content focused better on the key ideas related to the physical science topic area (the kinetic molecular theory) than the key ideas related to the life science (flow of matter and energy in ecosystems) and earth science (processes that shape the earth) topic areas. Hence, the programs received slightly higher ratings in the physical science topic area particularly on criteria that are sensitive to variations in content. In principle, programs might have scored better had different topic areas been selected. However, the specific ideas chosen for the study were important ones found in all materials examined; clearly, all the programs intended students to learn those ideas. Although it is possible that other ideas were targeted better, it seems unlikely that the authors would have done less than their best work on the key ideas examined in this study.

This study's findings are consistent with previous critiques of middle school programs which indicate that they cover many topics at a superficial level rather than treat a few topics in depth (Schmidt, McKnight, & Raizen, 1997; Tyson-Bernstein, 1988), focus on technical vocabulary and factual information rather than on main important ideas (Anderson & Roth, 1989), and do not attend to students' thinking and prior knowledge (Ciborowski, 1988; Eichinger & Roth, 1991). This study refines and extends these results by describing the programs' instructional shortcomings in terms of specific research-based criteria. In the present study, treating a topic in-depth is clarified to mean focusing on the key ideas within a topic, establishing a sense of purpose for learning the ideas, finding out where students are likely to be with respect to the ideas to know better how to focus instruction, helping students understand the ideas through appropriate representations unencumbered by needless details and terms, presenting sufficient relevant phenomena to make the ideas plausible, modeling how the ideas can be used to explain relevant phenomena, and providing sufficiently varied practice and reflection to promote transfer. None of the programs examined systematically aim such instructional strategies at any key idea. Even programs that explicitly claim to be based on a learning cycle do not aim an entire cycle on a key idea, although one or two stages of a cycle may do so. It is indeed noteworthy that conceptual development in such materials most of the time does not reflect the rationale of the learning cycle, i.e., to take each major concept through each stage of the cycle. Instead, different stages of the cycle focus on different concepts.

These findings support the value of a learning goals–based approach to the analysis of curriculum materials. Typically, analysis procedures examine separately the content and the instructional design of materials, considering, for instance, whether a material employs generally supportive instructional strategies such as hands-on activities and cooperative groups independent of what learning goals they may contribute to. In our approach, content and instructional design cannot be considered separately; criteria probe whether curriculum materials provide multiple and varied phenomena and guide student interpretation and reasoning about experiences with phenomena and readings for each key idea examined in the analysis. As our analyses show, instructional strategies may turn out to be aimed at only some key ideas or none at all. If one were to answer whether programs included a variety of hands-on activities in general, several programs might receive high ratings. However, by examining, for example, whether programs include a variety of experiences to support key ideas of the kinetic molecular theory, we often find that for some (or even all) key ideas there are hardly any supporting activities.

To increase reliable interpretations of criteria, our approach makes use of highly specific criteria, and explicit clarifications, indicators, and scoring guidelines. Nevertheless, individual biases from reviewers' philosophy about what constitutes good teaching and the role of curriculum materials, from experiences as teachers and researchers, and from expectations about quality of specific materials might still filter through the process. The team approach to the analysis of materials and consensus building of ratings in which team members from different backgrounds defend their own judgments and examine and comment on judgments and supporting evidence of others appears essential in reducing such individual bias.

Further issues related to the analysis procedure might constrain our findings and their use. The research literature provides guidance for making judgments about certain characteristics of curriculum materials but not for others. For example, the decisions on whether a problem that is used to frame a unit is likely to be interesting or motivating to middle school students had to rely mostly on the craft knowledge of experienced reviewers rather than on the research literature. Research-based curriculum materials in the physical and life science topic areas (for instance, Berkheimer, Anderson, & Blakeslee, 1988; Roth, 1997) provided topic-specific examples that illustrated high ratings of criteria. These examples provided a sound basis for the training of reviewers and the reconciliation of the teams' analyses in the physical and life science topic areas. Training of the earth science reviewers and reconciliations of the earth science analyses reports were constrained by the scarcity of research on students' learning about processes that shape the earth (and related topics) as well as by the lack of high-quality research-based materials that could serve as positive examples.

Finally, the Project 2061 curriculum-analysis criteria are intended to address features of curriculum materials that are most important for teaching and learning for the large majority of students and teachers. A particular teacher or set of students may not require a material to address every one of the criteria in the same way. For example, highly able students often can learn a concept with only a few examples or may need less guidance in interpreting the examples.

Implications

There is clearly a need for the development of new middle school science programs that better support teachers in their efforts to help students learn important ideas in science. The criteria used in this analysis, and the detailed findings pointing to current inadequacies in materials from this study could serve as guidelines in new materials development.

The development of new materials needs to focus on coherent sets of important ideas. *Benchmarks for Science Literacy* and *National Science Education Standards* can serve as sources for such ideas. The development of new materials needs also to incorporate findings from learning research, especially research on student thinking about particular topics in the science curriculum. This requires that relevant research findings be elaborated and communicated at a level of detail that makes them usable to curriculum developers and publishers of curriculum materials. For example, there is a need to explain the role of prior knowledge (both conceptual and cultural) in learning and its implications for design of learner-centered curricula. A relevant communication could clarify distinctions between incorrect or missing prior factual knowledge and deep-rooted beliefs that impede learning; explain the major beliefs that influence learning in selected topic areas; and describe topic-specific instructional strategies that build on or challenge these beliefs and have been found to be helpful in developing student thinking. Although there is much to be learned about the learning process and the role of curriculum materials in that process, it would be unfortunate if what we already know about student learning fails to make its way into the design of curriculum materials.

Although topic-specific research findings can provide the focus for learner-centered curriculum materials, this research base is incomplete. There are many topic areas in the science curriculum that have hardly been studied. However, even in areas in which research has been done, there is often a lack of research-based knowledge about which phenomena make ideas plausible to students, which representations are most effective in making ideas intelligible, and what kind of modeling will enable students to apply what they have learned to everyday situations (Heller, 2001). In these cases, curriculum materials development needs to include research activities. It may be impractical for publishers to conduct the in-depth studies of student learning and thinking that would be required. New partnerships between researchers and curriculum developers need to be developed so that teachers and students can have access to the best quality curriculum materials.

Finally, until better curriculum materials are developed, schools might consider keeping their current programs and focus their resources on professional development to help teachers supplement these programs. The procedure and detailed information in the reviews undertaken can be helpful to middle school science teachers as they start looking for ways to compensate for their textbooks' shortcomings (Kesidou, 2001).

Appendix A

Instructional Analysis Criteria

Project 2061's curriculum analysis procedure uses the following criteria, organized into seven categories, to determine the extent to which a material's instructional strategy is likely to support students learn the content. Each criterion in Categories I–VI is to be assessed with regard to specific learning goals, not just in general.

Category I: Identifying and Maintaining a Sense of Purpose

Conveying Unit Purpose. Does the material convey an overall sense of purpose and direction that is understandable and motivating to students?

Conveying Activity Purpose. Does the material convey the purpose of each activity and its relationship to others?

Justifying Activity Sequence. Does the material include a logical or strategic sequence of activities (versus just a collection of activities)?

Category II: Taking Account of Student Ideas

Attending to Prerequisite Knowledge and Skills. Does the material specify prerequisite knowledge or skills necessary to the learning of the key idea(s)?

Alerting Teacher to Commonly Held Student Ideas. Does the material alert teachers to commonly held student ideas (both troublesome and helpful)?

Assisting Teacher in Identifying Student Ideas. Does the material include suggestions for teachers to find out what their students think about familiar phenomena related to a key idea before the key ideas are introduced?

Addressing Commonly Held Ideas. Does the material explicitly address commonly held student ideas?

Category III: Engaging Students with Relevant Phenomena

Providing Variety of Phenomena. Does the material provide multiple and varied phenomena to support the key idea(s)?

Providing Vivid Experiences. Does the material include first-hand experiences with phenomena (when practical) and provide students with a vicarious sense of the phenomena when experiences are not first-hand?

Category IV: Developing and Using Scientific Ideas

Introducing Terms Meaningfully. Does the material introduce technical terms only in conjunction with experience with the idea or process and only as needed to facilitate thinking and promote effective communication?

Representing Ideas. Does the material include appropriate representations of the key ideas?

Demonstrating Use of Knowledge. Does the material demonstrate, model, or include suggestions for teachers on how to demonstrate or model skills or the use of knowledge?

Providing Practice. Does the material provide tasks or questions for students to practice skills or use of knowledge in a variety of situations?

Category V: Promoting Student Thinking about Phenomena, Experiences, and Knowledge

Encouraging Students to Explain Their Ideas. Does the material routinely include suggestions for having each student express, clarify, justify, and represent his or her ideas? Are suggestions made for when and how students will get feedback from peers and the teacher?

Guiding Student Interpretation and Reasoning. Does the material include tasks and/or question sequences to guide student interpretation and reasoning about experiences with phenomena and readings?

Encouraging Students to Think About What They Have Learned. Does the material suggest ways to have students check their own progress?

Category VI: Assessing Progress

Aligning to Goals. Assuming a content match between the curriculum material and the benchmark, are assessment items included that match the key ideas?

Testing for Understanding. Does the material assess understanding of key ideas and avoid allowing students a trivial way out, such as repeating a memorized term or phrase from the text without understanding?

Informing Instruction. Are some assessments embedded in the curriculum along the way, with advice as to how teachers might use the results to choose or modify activities?

Category VII: Enhancing the Science Learning Environment

Providing Teacher Content Support. Would the material help teachers improve their understanding of science, mathematics, and technology necessary for teaching the material?

Encouraging Curiosity and Questioning. Does the material help teachers to create a classroom environment that welcomes student curiosity, rewards creativity, encourages a spirit of healthy questioning, and avoids dogmatism?

Supporting All Students. Does the material help teachers to create a classroom community that encourages high expectations for all students, enables all students to experience success, and provides all students a feeling of belonging in the science classroom?

Appendix B

Indicators and Scoring Scheme for Criterion ‘‘Conveying Activity Purpose’’

Criterion: Does the material convey the purpose of each activity and its relationship to others?

Indicators of Meeting the Criterion

1. The material conveys or prompts teachers to convey the purpose of each activity to students.
2. The purpose is expressed in a way that is likely to be comprehensible to students.
3. The material encourages each student to think about the purpose of the activity.

4. The material conveys or prompts teachers to convey to students how each activity relates to the purpose of the unit.
5. The material periodically engages students in thinking about what they have learned so far and what they need to learn or do next.

Scoring Scheme

- Excellent (3): The material meets all 5 indicators.
- Satisfactory (2): The material meets any 3 of the 5 indicators.
- Poor (1): The material meets 1 of the 5 indicators.
- None (0): The material does not meet any of the indicators.

Appendix C

Programs Reviewed

- P1: McLaughlin, C., & Thompson, M. (1997). Glencoe physical science. New York: Glencoe/McGraw Hill.
  Daniel, L., Ortleb, E., & Bibbs, A. (1999). Glencoe life science. New York: Glencoe/McGraw Hill.
  Feather, R., & Snyder, S. (1997). Glencoe earth science. New York: Glencoe/McGraw Hill.
- P2: Atwater, M., Baptiste, H.P., Daniel, L., Hackett, J., Moyer, R., Takemoto, C., et al. (1995). Macmillan/McGraw Hill science. New York: Macmillan/McGraw Hill.
- P3: BSCS (1999–2000). BSCS middle school science & technology (Levels A, B, C). Dubuque, IA: Kendall/Hunt. [NSF-funded]
- P4: Maton, A., Hopkins, J., Johnson, S., LaHart, D., Warner, M., & Wright, J. (1997). Exploring physical/life/earth science. Upper Saddle River, NJ: Prentice Hall.
- P5: The Prime Science Education Group. (1998). Prime Science (Levels A, B, C). Dubuque, IA: Kendall/Hunt. [NSF-funded]
- P6: Decision Development Corporation. (1995). Science 2000. Lexington, MA: D. C. Heath and Company. [NSF-funded, multimedia]
- P7: DiSpezio, M.A., Linner-Luebe, M.S., Lisowski, M., Skoog, G., & Sparks, B. (1997). Science insights. Exploring matter and energy. Menlo Park, CA: Addison-Wesley.
  DiSpezio, M.A., Linner-Luebe, M.S., Lisowski, M., Skoog, G., & Sparks, B. (1997). Science insights. Exploring living things. Menlo Park, CA: Addison-Wesley.
  DiSpezio, M.A., Linner-Luebe, M.S., Lisowski, M., Skoog, G., & Sparks, B. (1997). Science insights. Exploring earth and space. Menlo Park, CA: Addison-Wesley.
- P8: Aldridge, B., Aiuto, R., Ballinger, J., Barefoot, A., Crow, L., & Feather, R., Jr., et al. (1998). Science interactions (Courses 1, 2, 3). New York: Glencoe/McGraw Hill.
- P9: Morrison, E., Moore, A., Armour, N., Hammond, A., Haysom, J., Nicoll, E., et al. (1997). Science plus (Levels Red, Green, Blue). Austin, TX: Holt, Rinehart and Winston.

References

Albanese, A., & Vicentini, M. (1997). Why do we believe that an atom is colorless? Reflections about the teaching of the particle model. Science and Education, 6, 251–261.

Aldridge, B., Aiuto, R., Ballinger, J., Barefoot, A., Crow, L., & Feather, R., Jr., et al. (1998). Science interactions. Teacher wraparound edition (Course 3). New York: Glencoe/McGraw Hill.

Alexander, P., & Kulikowich, J. (1994). Learning from physics text: A synthesis of recent research. Journal of Research in Science Teaching, 31, 895–911.

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

Anderson, C.W. (1991). Policy implications of research on science teaching and teacher's knowledge. In Kennedy, M., (Ed.), Teaching academic subjects to diverse learners (pp. 5–30). New York: Teachers College Press.

Anderson, C.W., & Roth, K.J. (1989). Teaching for meaningful and self-regulated learning of science. In Brophy, J., (Ed.), Advances in Research on Teaching: Vol. 1 (pp. 265–310). Greenwich, CT: JAI Press.

Anderson, C.W., & Smith, E.L. (1984). Children's preconceptions and content-area textbooks. In Duffy, G.G., Rochler, L.R., & Mason, J. (Eds.), Comprehension instruction: Perspectives and suggestions (pp. 187–220). New York: Forgman.

Anderson, C.W., & Smith, E. (1987). Teaching science. In Richardson-Koehler, V. (Ed.), The educator's handbook: A research perspective (pp. 84–111). New York: Longman.

Andersson, B. (1990). Pupils' conceptions of matter and its transformations (age 12–16). In Lijnse, P.L., Licht, P., de Vos, W., & Waarlo, A.J. (Eds.), Relating macroscopic phenomena to microscopic particles (pp. 12–35). Utrecht, The Netherlands: Centre for Science and Mathematics Education.

Arons, A. (1990). A guide to introductory physics teaching. New York: Wiley.

Atwater, M., Baptiste, H.P., Daniel, L., Hackett, J., Moyer, R., Takemoto, C., et al. (1995). Macmillan/McGraw Hill science. Changes in matter. Teacher's planning guide. New York: Macmillan/McGraw Hill.

Ball, D.L., & Cohen, D.K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? Educational Researcher, 25, 6–8, 14.

Ball, D.L., & Feiman-Nemser, S. (1988). Using textbooks and teachers' guides: A dilemma for beginning teachers and teacher educators. Curriculum Inquiry, 18, 401–423.

Bazler, J., & Simonis, D. (1991). Are high school chemistry textbooks gender fair? Journal of Research in Science Teaching, 28, 353–362.

Berkheimer, G., Anderson, C., & Blakeslee, T. (1988). Matter and molecules. East Lansing: Michigan State University, Institute for Research on Teaching [On-line]. Available: http://ed-web3.educ.msu.edu/reports/matter-molecules/default.htm. Retrieved January 14, 2001.

Bishop, B., & Anderson, C. (1990). Student conceptions of natural selection and its role in evolution. Journal of Research in Science Teaching, 27, 415–427.

Blumenfeld, P., Soloway, E., Marx, J., Krajcik, J., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. Educational Psychologist, 26, 369–398.

Boulanger, F.D. (1981). Instruction and science learning: A quantitative synthesis. Journal of Research in Science Teaching, 18, 311–327.

Brook, A., Briggs, H., & Driver, R. (1984). Aspects of secondary students' understanding of the particulate nature of matter. Leeds, UK: University of Leeds, Center for Studies in Science and Mathematics Education.

Brown, D., & Clement, J. (1992). Classroom teaching experiments in mechanics. In Duit, R., Goldberg, F., & Niedderer, H. (Eds.), Research in physics learning: Theoretical issues and

empirical studies (pp. 380–397). Kiel, Germany: Institute for Science Education at the University of Kiel.

BSCS (2000). BSCS middle school science & technology. Teacher's edition (Level B). Dubuque, IA: Kendall/Hunt.

Caldwell, A., & Stern, L. (2000, April). Can middle-school science textbooks help students learn important ideas? Findings from Project 2061's curriculum evaluation study: Earth science. Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Champagne, A., Gunstone, R., & Klopfer, L. (1985). Instructional consequences of students' knowledge about physical phenomena. In West, L., & Pines, A.L. (Eds.), Cognitive structure and conceptual change (pp. 61–90). Orlando, FL: Academic Press.

Chi, M. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In Glaser, R. (Ed.), Advances in instructional psychology (pp. 161–238). Mahway, NJ: Erlbaum.

Chiapetta, E., Sethna, G., & Fillman, D. (1993). Do middle school life science textbooks provide a balance of scientific literacy themes? Journal of Research in Science Teaching, 30, 787–797.

Chinn, C., & Brewer, W. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. Review of Educational Research, 63, 1–49.

Ciborowski, J. (1988). Improving textbook usability (final report). Newton, MA: Education Development Center. (ERIC Document Reproduction Service No. ED 303 934)

DiSpezio, M.A., Linner-Luebe, M.S., Lisowski, M., Skoog, G., & Sparks, B. (1997). Science insights. Exploring matter and energy. Teacher's edition. Menlo Park, CA: Addison-Wesley.

Driver, R. (1983). The pupil as scientist? Milton Keynes: Open University Press.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23, 5–12.

Eaton, J.F., Anderson, C.W., & Smith, E.L. (1984). Student preconceptions interfere with learning: Case studies of fifth-grade students. Elementary School Journal, 64, 365–379.

Education Market Research. (1998). Elementary and middle school science market: Teaching methods, materials used and needed, and market size. Rockaway Park, NY: Author.

Eichinger, D., & Roth, K. (1991). Critical analysis of an elementary science curriculum: Bouncing around or connectedness? East Lansing, MI: Center for the Learning and Teaching of Elementary Subjects. (ERIC Document Reproduction Service No. ED 340 611)

Eltinge, E., & Roberts, C. (1993). Linguistic content analysis: A method to measure science as inquiry in textbooks. Journal of Research in Science Teaching, 30, 65–83.

Ericsson, K.A., Krampe, R.T., & Tesche-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100, 363–406.

Good, R. (1993). Science textbook analysis. Journal of Research in Science Teaching, 30, 619.

Good, R., & Shymansky, J. (1986). Issues regarding the establishment of criteria for the analysis and selection of science textbooks. Paper presented at a joint meeting of the School Division–Association of American Publishers, the National Association for Research in Science Teaching, and the National Council of Teachers of Mathematics, Atlanta, Georgia. (ERIC Reproduction Service No. ED 267 987)

Guzzetti, B., Snyder, T., Glass, G., & Gamas, W. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. Reading Research Quarterly, 28, 117–159.

Heller, P. (2001, February). Lessons learned in the CIPS (Constructing Ideas in Physical Science) Curriculum Project. Paper presented at the AAAS Project 2061 Science Textbook Conference, Developing Textbooks That Promote Science Literacy, Washington, DC.

Hesse, J.J., & Anderson, C.W. (1992). Students' conceptions of chemical change. Journal of Research in Science Teaching, 29, 277–299.

Holliday, W. (1990). Textbook illustrations. Fact or filler? Science Teacher, 57, 27–29.

Holliday, W., Yore, L., & Alvermann, D. (1994). The reading-science learning-writing connection: Breakthroughs, barriers, and promises. Journal for Research in Science Teaching, 31, 877–893.

Hollon, R.E., Roth, K.J., & Anderson, C.W. (1991). Science teachers' conceptions of teaching and learning. In Brophy, J. (Ed.), Advances in research on teaching: Volume 2. Teacher's knowledge of subject matter as it relates to their teaching practice (pp. 145–186). Greenwich, CT: JAI Press.

Hynd, C., McWhorter, J., Phares, V., & Suttles, W. (1994). The role of instructional variables in conceptual change in high school physics topics. Journal of Research in Science Teaching, 31, 933–946.

Jimenez, M.P.A. (1994). Teaching evolution and natural selection: A look at textbooks and teachers. Journal of Research in Science Teaching, 31, 519–535.

Kesidou, S. (2001). Aligning curriculum materials with National Science Standards: The role of Project 2061's curriculum-analysis procedure in professional development. Journal of Science Teacher Education, 12, 47–65.

Lee, O., Eichinger, D., Anderson, C.W., Berkheimer, G.D., & Blakeslee, T.D. (1993). Changing middle school students' conceptions of matter and molecules. Journal of Research in Science Teaching, 30, 249–270.

Linn, M., & Burbules, N. (1993). Construction of knowledge and group learning. In Tobin, K. (Ed.), Practice of constructivism in science education (pp. 91–119). Washington, DC: AAAS.

Lynch, S. (2000). Equity and science education reform. Mahwah, NJ: Erlbaum.

Malone, T.W., & Lepper, M.R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In Snow, R., & Farr, M. (Eds.), Aptitude, learning, and instruction: Cognitive and affective process analyses Vol. 3 (pp. 223–253). Hillsdale, NJ: LBJ.

Maton, A., Hopkins, J., Johnson, S., LaHart, D., Warner, M., & Wright, J. (1997). Exploring physical science. Teacher's edition. Upper Saddle River, NJ: Prentice Hall.

McDermott, L. (1991). Millican lecture 1990. What we teach and what is learned—closing the gap. American Journal of Physics, 59, 301–315.

McLaughlin, C., & Thompson, M. (1997). Glencoe physical science. Teacher wraparound edition. New York: Glencoe/McGraw Hill.

McNamara, T.P., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. Cognition and Instruction, 14, 1–43.

Minstrell, J. (1984). Teaching for the understanding of ideas: Forces on moving objects. In Anderson, C.W. (Ed.), Observing science classrooms: Perspectives from research and practice. 1984 Yearbook of the Association for the Education of Teachers in Science (pp. 55–73). Columbus, OH: ERIC Center for Science, Mathematics and Environmental Education. (ERIC Document Reproduction Service No. ED 255 355)

National Educational Goals Panel. (1994). Data volume for the National Education Goals Report (Vol. 1). Washington, DC: Author.

National Research Council. (1994). Learning, remembering, believing: Enhancing human performance. Washington, DC: National Academy Press.

National Research Council. (1996). National science education standards. Washington, DC: National Academy Press.

National Research Council. (1999). Improving student learning. A strategic plan for education research and its utilization. Washington, DC: National Academy Press.

National Research Council. (2000a). How people learn. Brain, mind, experience, and school. Washington, DC: National Academy Press.

National Research Council. (2000b). How people learn. Bridging research and practice. Washington, DC: National Academy Press.

Nussbaum, J. (1985). The particulate nature of matter in the gaseous phase. In Driver, R., Guesne, E., & Tiberghien, A. (Eds.), Children's ideas in science (pp. 124–144). Milton Keynes, UK: Open University Press.

Pivard, L.P., & Straw, S.B. (2000). The effect of talk and writing on learning science: An exploratory study. Science Education, 84, 566–593.

Potter, E., & Rosser, S. (1992). Factors in life science textbooks that may deter girls' interest in science. Journal of Research in Science Teaching, 29, 669–686.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. Journal of the Learning Sciences, 2, 235–276.

Roseman, J., Kesidou, S., & Stern, L. (1996, November). Identifying curriculum materials for science literacy: A Project 2061 evaluation tool. Paper presented at the National Research Council Colloquium, ''Using the National Science Education Standards to Guide the Evaluation, Selection, and Adaptation of Instructional Materials,'' Washington, DC.

Roth, K.J. (1984). Using classroom observations to improve science teaching and curriculum materials. In Anderson, C.W. (Ed.), Observing science classrooms: Perspectives from research and practice. 1984 yearbook of the Association for the Education of Teachers in Science (pp. 77–102). Columbus, OH: ERIC/SMEAC.

Roth, K. (1991). Reading science texts for conceptual change. In Santa, C.M., & Alverman, D.E. (Eds.), Science learning processes and applications. Newark, DE: International Reading Association.

Roth, K. (1997). Food for plants: Student text and teacher's guide. East Lansing, MI: Michigan State University.

Roth, K. (2001, February). Student-focused curriculum materials development: The case of ''Food for Plants.'' Paper presented at the AAAS Project 2061 Science Textbook Conference, ''Developing textbooks that promote science literacy,'' Washington, DC.

Salomon, G., & Perkins, D. (1998). Individual and social aspects of learning. Review of Research in Education, 23, 1–24.

Scott, P.H., Asoko, H.M., & Driver, R.H. (1992). Teaching for conceptual change: A review of strategies. In Duit, R., Goldberg, F., & Niedderer, H., (Eds.), Research in physics learning: Theoretical issues and empirical studies (pp. 310–329). Kiel, Germany: IPN.

Schmidt, W., McKnight, C., & Raizen, S. (1997). Executive summary. A splintered vision: An investigation of U.S. science and mathematics education. Lansing, MI: U.S. National Research Center, Michigan State University.

Shiland, T. (1997). Quantum mechanics and conceptual change in high school chemistry textbooks. Journal of Research in Science Teaching, 34, 535–545.

Smith, E., & Anderson, C.W. (1983, April). The effects of teacher's guides on teacher planning and classroom instruction in activity-based science. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Smith, E.L., & Anderson, C.W. (1984). Plants as producers: A case study of elementary school science teaching. Journal of Research in Science Teaching, 21, 685–698.

Smith, E.L., Blakeslee, T., & Anderson, C.W. (1993). Teaching strategies associated with conceptual change learning in science. Journal of Research in Science Teaching, 30, 111–126.

Songer, N.B., & Linn, M.C. (1991). How do students' views of science influence knowledge integration? Journal of Research in Science Teaching, 28, 761–784.

Spiro, R.J., Feltovich, P.J., Coulson, R.L., & Anderson, D.K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconception in advanced knowledge acquisition. In Vosniadou, S. & Ortony, A. (Eds.), Similarity and analogical reasoning (pp. 498–531). Cambridge, MA: Cambridge University Press.

Stern, L. & Roseman, J. (2000, April). Can middle-school science textbooks help students learn important ideas? Findings from Project 2061's curriculum evaluation study: Life science. Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Strike, K. & Posner, G.S. (1985). A conceptual change view of learning and understanding. In West, L. & Pines, A.L. (Eds.), Cognitive structure and conceptual change (pp. 211–231). Orlando, FL: Academic Press.

Tasker, R., & Freyberg, P. (1985). Facing the mismatches in the classroom. In Osborne, R. & Freyberg, P. (Eds.), Learning in science: The implications of children's science (pp. 66–80). Auckland: Heinemann.

The Prime Science Education Group. (1998). Prime science. Level C (student book). Dubuque, IA: Kendall/Hunt.

Treagust, D., & Harrison, A. (2000). In search of explanatory frameworks: An analysis of Richard Feynman's lecture "Atoms in Motion." International Journal of Science Education, 22, 1157–1170.

Tyson-Bernstein, H. (1988). America's textbook fiasco: A conspiracy of good intentions. Washington, DC: Council for Basic Education.

Vachon, M., & Haney, R. (1991). A procedure for determining the level of abstraction of science reading material. Journal of Research in Science Teaching, 28, 343–352.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. Cognition and Instruction, 16, 3–118.

Wise, K.C, & Okey, J.R. (1983). A meta-analysis of the effects of various science teaching strategies on achievement. Journal of Research in Science Teaching, 20, 419–435.

Woloshyn, V.E., Willoughby, T., Wood, E., & Pressley, M. (1990). Elaborative interrogation facilitates adult learning of factual paragraphs. Journal of Educational Psychology, 82, 513–524.