

## **Part 3**

What Issues Arise  
with Inquiry Learning  
and Teaching?



# **Instructional, Curricular, and Technological Supports for Inquiry in Science Classrooms<sup>1</sup>**

Joseph Krajcik, Phyllis Blumenfeld,  
Ron Marx, and Elliot Soloway

---

New approaches to science instruction feature inquiry on the part of students as essential for student learning (Lunetta, 1998; Roth, 1995). The assumption is that students need opportunities to find solutions to real problems by asking and refining questions, designing and conducting investigations, gathering and analyzing information and data, making interpretations, drawing conclusions, and reporting findings. In the spirit with recommendations by the American Association for the Advancement of Science (AAAS) (1993), the National Research Council (NRC) (1996) argues that “there needs to be a de-emphasis on didactic instruction focusing on memorizing decontextualized scientific facts, and there needs to be new emphasis placed on inquiry-based learning focusing on having students develop a deep understanding of science embedded in the everyday world.”

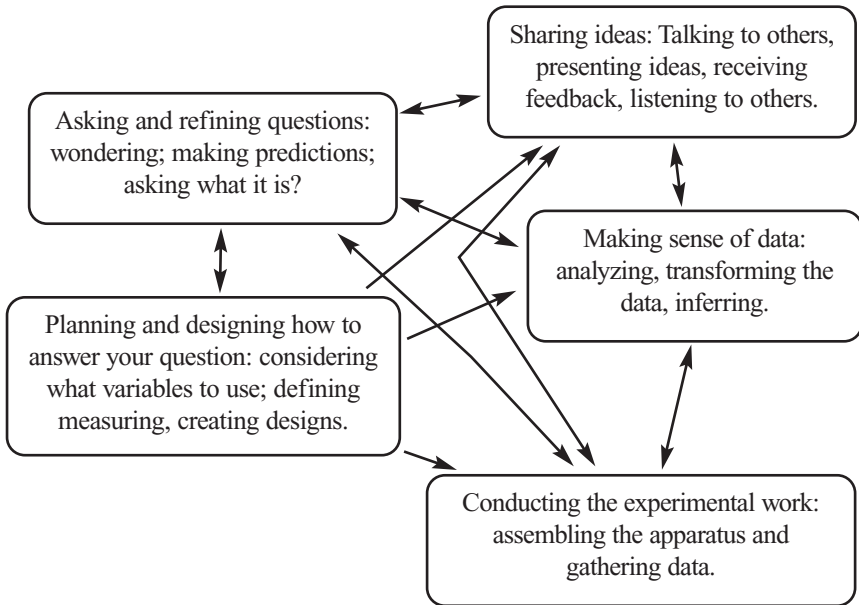
Evidence indicates that engagement in inquiry can bring students to deeper understanding of science content and processes (e.g., Brown & Campione, 1994; Cognition and Technology Group at Vanderbilt, 1992; Metz, 1995). But our work (Krajcik et al., 1998), along with that of others (Brown & Campione, 1994; Linn, 1998; Roth, 1995), has demonstrated that the cognitive demands

that inquiry places on learners require considerable support. Students need help to become knowledgeable about content, skilled in using inquiry strategies, proficient at using technological tools, productive in collaborating with others, competent in exercising self-regulation, and motivated to sustain careful and thoughtful work over time. Describing problems students encounter as they engage in inquiry and finding ways to ameliorate those problems have received considerable attention recently (Hmelo & Williams, 1998; McGilly, 1994; Blumenfeld et al., 1998). In this paper, we describe inquiry in more detail, discuss ways to aid students via instructional, curriculum, and technological supports, and then illustrate how these have been applied to specific phases on inquiry where students encounter difficulties.

### WHAT IS INQUIRY AND WHY USE IT?

Broadly conceived, inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work (NRC, 1996). Inquiry is not a linear process. Phases interact; preliminary findings, for instance, may result in a decision to revise the original question or to alter data collection procedures. Figure 1 shows a model of inquiry (Krajcik et al., 1998).

Renewed interest in inquiry comes from research that shows that students' understanding of scientific ideas and scientific process is limited, so that many who do well on tests cannot apply their knowledge outside the classroom. New approaches to instruction assign primary importance to the way in which students make sense of what they are learning, rather than to how teachers deliver information. The assumption of such constructivist programs (Fensham, Gunstone, & White, 1994) is that integrated and usable knowledge develops when learners create multiple representations of ideas and are engaged in activities that require them to use that knowledge. Inquiry promotes development, transformation, and representation of ideas and helps learners understand how knowledge is generated in different disciplines. The emphasis is on depth, not breadth. In addition, conversation with others is an important way for students to exchange information, explain, and clarify their ideas, consider others' ideas, and expand their understanding.

**FIGURE 1: THE INVESTIGATIVE WEB**

Among approaches that use inquiry, which include authentic tasks, artifacts, alternative assessments, technological tools, and collaboration, our work has students pursue investigations to answer a driving question related to their everyday experience. In finding answers to the question students learn scientific concepts, engage in scientific processes, and gain a better understanding of the discipline (Blumenfeld et al., 1991; Krajcik et al., 1994). Others rely on anchored experiences created on videotape (Sherwood, Petrosino, & Lin, 1998). Constructivist approaches often emphasize the production of artifacts such as multimedia documents, models, presentations, or demonstrations. Producing such artifacts allows students to apply information and represent knowledge in a variety of ways. Artifacts serve as a way for teachers to see how students are thinking, and for students to share their ideas and receive feedback which can be incorporated into revisions. Rather than rely on standardized tests, which have been criticized for concentrating on knowledge of isolated facts, the use of alternative assessments that have some value beyond the classroom is encouraged (for example, Newmann & Archibald, 1992; Perkins, 1992). These assessments, such as public performances, creation of museum

exhibits or reports to local groups, require students to exhibit mastery of the discipline and to integrate prior knowledge with new.

Tool use is another core element of these approaches. Recent interest has centered on the use of learning technologies, such as interactive video technology (compact discs or videodisks), telecommunications, microcomputer-based laboratories, modeling, and the World Wide Web. Learning technologies can help learners solve complex and ambiguous problems by providing access to information and opportunities to collaborate, investigate, and create artifacts. Tools can extend and amplify learners' thinking because they reduce the cognitive load for students, moving from students to the computer some routine tasks like calculating, creating graphs, or depicting data in different forms (Salomon, Perkins, & Globerson, 1991).

Collaboration and conversation also are stressed. As students engage in conversation, they draw on the knowledge and expertise of others, reflect on their own ideas, and internalize modes of knowledge and thinking represented and practiced in the subject (Bruer, 1994).

## SUPPORTS FOR INQUIRY

Supports to students in the inquiry process include instructional, curricular, and technological aids. These can work independently as well as in conjunction.

### Instructional Supports

During inquiry the teacher serves as a learner as well as a guide or facilitator. Benchmark lessons (Krajcik, Czerniak, & Berger, 1999) introduce students to relevant content and skills before and during inquiry. The teacher helps students develop the thinking strategies used by experts, like heuristics for generating questions or interpreting data. They also help students become more metacognitive, attentive to planning, monitoring work, and evaluating their progress.

**Scaffolding.** For the teaching and learning situation, Collins, Brown, and Newman (1989) use the analogy of a cognitive apprenticeship. The teacher scaffolds instruction by breaking down tasks, using modeling and coaching to teach strategies for thinking, provides feedback that helps students diagnose their problems, and gradually releases responsibility to learners to perform these functions on their own. The emphasis is on helping students to become more like experts in their thinking about generating questions, using strategies to design inquiries to find solutions to questions, and evaluating the results of their

efforts by mirroring heuristics and stratagems that experts have been found to use. These types of scaffolds can be used during each phase of inquiry.

Krajcik, Czerniak, and Berger (1999) offer these definitions and examples of scaffolds.

- ▶ Modeling is the process by which a more knowledgeable individual illustrates to the learner how to do or think about a task. For example, a teacher could demonstrate how to use the concept of “mean” to analyze data or how to read a pH meter. Many science processes can be modeled for students. Some of these include illustrating for students how to ask questions, plan and design investigations, or form conclusions.
- ▶ Coaching involves providing suggestions and asking questions to help the student improve knowledge or skills. For example, a teacher could make suggestions to a student about how to make more precise measurements when reading a spring scale. Other forms of coaching can include asking thought provoking questions (such as “How do your data support your conclusion?”), giving students sentence stems (for example, “My data supports my conclusion because...”), and supplying intellectual or cognitive prompts (such as asking students to write down predictions, give reasons, and elaborate answers).
- ▶ Sequencing is breaking down a larger task into step-by-step sub-tasks so a learner can focus on completing just one sub-task at a time rather than the entire task at once. For instance, the teacher might break down the process of investigations into various components and not allow the learner to proceed to the next step until completing the previous step. For example, the teacher could require the learner to complete a plan before moving on to building an apparatus.
- ▶ Reducing complexity involves hiding complex understandings or tasks until the learner has mastered simpler understandings or sub-tasks. The classical example here is helping a child learn to ride a bicycle by using training wheels. In science classrooms, this might mean a teacher uses an analogy to reduce the complexity of a concept. For instance, the teacher could compare DNA to the instructions for building a model airplane.
- ▶ Highlighting the critical features of a concept or task is another way a knowledgeable other can support the learning of another person. A teacher could point out to young students that animals called mammals

all have hair—hair is a key feature. As another example, in teaching a student how to focus a microscope, the teacher might point out that a basic step in focusing the object on the slide is to start always with the lowest powered lens first.

- ▮ Using visual tools can help students understand a concept or task (Hyerle, 1996; Parks & Black, 1992). Visual tools are pictorial prompts that help students understand their own thinking process. Visual tools also help make abstract ideas more concrete by organizing ideas or illustrating relationships. To develop understandings of kinetic molecular theory, students could use a computer simulation that represents the particle nature of matter.

Prediction, Observation, and Explanation Cycles. Another frequently used technique that promotes linking prior and new knowledge is the cycle of prediction, observation, and explanation (POE) (White & Gunstone, 1992). Students are asked to draw on prior knowledge to make predictions about what will occur during a demonstration, what they might find when searching for information, or what the results of an experiment might be. They can make individual predictions, share them with a group, discuss reasons for their predictions and come to some consensus about what might occur as they exchange ideas. Teachers can also make a prediction, thinking aloud to model how they draw on what they know to determine what might happen, or coach students as they consider possibilities, pointing out things to consider. Next students observe the phenomena and record their observations. As in the prediction phase, students can make individual observations, share them with a group, and come to some consensus about what they observed. Teachers provide scaffolds by demonstrating how they record and organize data, how students may reduce complexity of complicated observations by creating charts, or how teachers insure that the data are complete and correct. Finally, students compare their predictions with the observations and develop explanations about inconsistencies. Here again teachers can model how they generate explanations and consider whether the explanation is adequate, coach students as they develop explanations, and highlight various essential features to consider. The point is to emphasize justification and exchange of ideas.

Concept Maps. Concept maps are visual representations of the relationship among ideas. The maps are organized hierarchically with most important and inclusive concepts at the top. Related ideas are clustered around the overarching concepts and are linked. Maps are judged on the accuracy of the hierarchy and linking of ideas. Such mapping helps students organize, structure, and

connect information and results in more meaningful understanding of ideas. Mapping also aids in the retrieval and facilitates the transfer of ideas.

Novak and Gowin in 1984 created the concept map as a tool to assess changes in learning. But concept mapping can be used as well to elicit student understandings prior to exploring a question. They are also an excellent way for students to track concepts that are being explored during inquiry. As the investigation continues, students make new concept maps that integrate new information with previous understandings. Comparing earlier and later versions of their maps can show students how their conceptual understanding is developing. Another useful approach is to have students compare their concepts maps with those of other students; to discuss and resolve differences. Conversational aids developed by Coleman (1998) to improve discussion in small groups of ideas and quality of explanations during the construction of concept maps are described in the section on collaboration.

Writing. Writing is another way to enhance student understanding. As students write they must retrieve, synthesize, and organize information. Production of a written document requires learners to clarify their thoughts (Santa & Havens, 1991) and also provides teachers with a window on student thinking. Keys (1994) shows that using collaborative writing guided by a series of prompts during the POE cycle improved the student's ability to draw conclusions, formulate models, and compose explanations that synthesized prior knowledge, observations, and other sources of information.

As in other fields, journals are one form of writing that is receiving considerable attention in science education (e. g., Audet, Hickman, & Dobrynina, 1996; Britsch & Shepardson, 1997). Bass and Baxter in 1998 studied how fifth-grade teachers made use of such notebooks, such as to write and draw, as a way to take notes, record observations, practice science skills, and summarize information and as a resource for teachers to monitor completion of tasks, assess understanding of specific concepts, and provide feedback. They found that rather than use the writing as a way for students to work through and demonstrate their understanding, teachers often controlled the writing, dictating what should be included and how. For instance, the authors determined that the notebooks were used to record procedure and findings rather than to explain conclusions and reasons for them and underrepresented the types of conversations students had about strengths and weaknesses of different methods and about the meaning of their data. The authors argue that for notebooks to fulfill their potential students should be asked to record their decisions and explain their thinking, keep track of their ideas and points made in conversations, and in general, center on substance rather than procedures.

### Design of Curriculum Materials to Support Inquiry

In response to educational recommendations, many new curricular packages have been designed to promote inquiry, which also incorporate technology. Examples include *Scientists in Action* developed by the Cognition and Technology Group at Vanderbilt (Sherwood et al., 1998), Linn's (1998) *Computers as Learning Partners*, Songer's (1993) *Kids as Global Scientists*, and Edelson's (1998) *WorldWatcher*. Evidence indicates that these approaches help students achieve deeper understanding. Under the auspices of the Center for Learning Technologies in Urban Schools, The University of Michigan in collaboration with Detroit Public Schools is developing year-long curriculum materials for middle school students. The design principles<sup>2</sup> underlying these materials incorporate learning theory, our own experience, and the experience and suggestions of teachers and professional educators. The curriculum materials are organized into projects that promote understanding of science concepts via inquiry, are predicated on constructivist principles (Marx et al., 1998), and address the needs of diverse students (Atwater, 1994; Ladson-Billings, 1995). These design principles are the basis for curriculum materials.

1. **Standards.** Materials are designed to meet school district curriculum guidelines, which are congruent with AAAS' Benchmarks (1993) and the NRC's *Standards* (1996).
2. **Contextualization.** A "driving question" that draws on students' experiences gives context to scientific ideas and makes inquiry authentic. In the process of exploring answers to the question, students encounter and come to understand these scientific ideas. The question must therefore encompass rich scientific content so that it is intellectually worthwhile. It is chosen with the advice of teachers, parents, and content experts. For instance, students study chemistry by investigating the question, "Why does our air smell bad and is it bad for us?"
3. **Anchoring.** Students begin exploring the question via a common experience they can refer back to during the course of the project. These experiences, such as collecting and analyzing samples of water from the local river, help to anchor the question (CTGV, 1992).
4. **Inquiry.** In exploring the driving question, students raise questions, design investigations, apparatus, and procedures for collecting data, gather and analyze data, and present results (Krajcik et al., 1998).

5. **Technology tools.** Each project is designed to incorporate technology tools that are most appropriate for finding solutions to the question. In the project “Why do I need to wear a bicycle helmet?” students use motion probes to explore distance-time graphs and velocity-time graphs. In “What is the quality of water in our river?” students use Model-It to create relations among various factors affecting water quality and use probes to monitor the water.
6. **Collaboration.** Students work with peers and with others outside the classroom: community members, university students, and students in other schools.
7. **Community involvement.** The questions on which students work mesh science with issues such as environmental quality and disease that are likely to be of interest to community organizations and to the family. Community organizations serve as sources of information about local problems and local expertise with respect to the question under study, as sites where students access technology after school, and as audiences for student work.
8. **Scaffolding.** The curriculum materials are scaffolded within projects so that students are introduced to concepts and to science processes in a manner that guides their learning. The emphasis is on modeling of skills and heuristics, such as how to evaluate the quality of a question, how to create charts to keep track of data collection or how to represent data in different ways. The teacher, the structure of the tasks, and the technology provide scaffolds within a project. Teachers are given suggestions about when to model, coach, give feedback, and present benchmark lessons. Tasks are structured to reduce complexity so that certain concepts or inquiry strategies are highlighted and questions that foster thoughtfulness provided. Technology scaffolds students by providing multiple representations, hiding complexity, and ordering and guiding processes such as planning, building, and evaluating.
9. **Sequencing.** The curriculum materials also provide support for students by sequencing inquiry processes and scientific concepts. Early in the middle school years, projects are structured tightly to minimize complexity. Tasks are chosen to illustrate particular inquiry strategies and the enabling power of technologies. This tight structuring affords students the opportunity to experience all phases of the inquiry process and to build an understanding of how all the phases fit

together. Later students are given more responsibilities for designing and conducting investigations on their own. Projects also are sequenced so that throughout the middle school years concepts are revisited. As a result, students develop rich understandings of how ideas are related to one another and to different scientific phenomena.

10. **Development of artifacts.** Throughout the projects students create a variety of artifacts such as investigative designs, plans for data collection, laboratory notebooks and models that both represent and help build understandings. These artifacts serve as embedded assessments by the teacher. Also, they can be shared, critiqued, and revised to enhance understanding. Students also create final artifacts such as oral or written presentations or multimedia documents that are exchanged with classmates, and with others in the school and the community. Having students demonstrate their learning in ways that go beyond the classroom is one feature of authentic instruction (Newman & Welage, 1993). Detailed rubrics assist teachers in evaluating artifacts to gauge student understanding.

### An Example Project

In a project on motion and force students explore the driving question “Why do I have to wear a bike helmet when I ride my bike?” During this eight-week unit, designed for eighth graders, students inquire into the physics of collision. It begins with a dramatic short videotape illustrating how bike accidents can result in brain injury. Then comes a series of demonstrations using an unprotected egg riding a cart, representing a student riding a bicycle, to illustrate the possible results of a collision. This demonstration is revisited periodically throughout the project and serves as the anchoring experience that students return to as they explore concepts of inertia, velocity, acceleration, force, and the relationships among them. It is also the focus of the final artifact; students design a helmet to protect the egg during a collision.

While exploring aspects of the driving question, students participate in several investigations supported by technology. They design experiments to examine the relationship between mass and inertia. Students study velocity and acceleration by collecting real time data using motion probes, which allows them to see these data immediately on the computer screen. They also learn how to read and interpret motion graphs. An investigation of gravity and mass involves collecting and interpreting information with the use of photogates to determine

velocity. Students use motion probes again in designing and testing their egg helmets. These designs and the results of the testing are presented and discussed with the class. We also encourage the teacher to invite visitors from local safety and community organizations who attend the presentations.

### Technology Design to Support Inquiry

Although inquiry can be done in classrooms without the aid of technology, learning technologies expand the range of questions that can be investigated, the types of information that can be collected, the kinds of data representations that can be displayed, and the products that that students can create to demonstrate their understandings. Such tools enable students to gather information about their questions on the World Wide Web, collect real time data using probes and other portable technologies, make models, graphs, and tables as a means of visually displaying data and quickly comparing different results, and illustrate their understandings in a variety of ways (for example, multimedia presentations). Students can work collaboratively with others in and outside the classroom. Examples of these tools are Knowledge Integration Environment developed at the University of California, Berkeley (Linn, 1998), and Worldwatcher, developed at Northwestern University (Edelson, 1998). The systems are integrated, designed to promote different fields of inquiry and allow for sharing. The tools are not specific to any particular content and they can be used to solve a range of problems and concepts. Because they can be used in different science classes across different grades, students can become proficient users of the tools and knowledgeable about the process of inquiry they support. The Investigators' Workshop is an example of learning technologies developed at the University of Michigan.

The Investigators' Workshop. The Investigators' Workshop,<sup>3</sup> is a suite of computational tools, based on learner-centered design (see next section), developed to enable sustained inquiry (Soloway & Krajcik, 1996). As described in Table 1, the tools support data collection, data visualization and analysis, dynamic modeling, planning, information gathering from the University of Michigan digital library and the Internet and web publishing (Jackson et al., 1996; Soloway, 1997; Soloway & Krajcik, 1996; Spitulnik et al., 1997). These tools have been revised several times in response to how students use them, the supports needed, and the types of artifacts produced by students.

**TABLE 1. THE INVESTIGATORS' WORKSHOP: SCAFFOLDED TOOLS FOR LEARNERS ENGAGED IN SUSTAINED SCIENCE INQUIRY**

NAME	FUNCTION	INQUIRY SUPPORT
Artemis	Supports on-line search and information gathering and evaluation using the UM Digital Library and Internet	Information gathering
Middle Years Digital Library Website	Provides support to students and teachers for carrying out on-line search activities to support inquiry	Information gathering and evaluating
Portable Computers and Probes	Microcomputer-based laboratories for portable computers; allows students to collect experimental data outside classroom by connecting various probes to the serial port	Data gathering
RiverBank	Water quality database tied to GREEN's field guide to water monitoring	Data sharing and storage
DataViz	Data visualization tool; supports students as they strive to see relationships and patterns in data both self-collected and gathered from on-line sources using visualization and analysis techniques	Data visualization and analysis

EChem	3-D molecular visualization tool; allows students to build three dimensional representations of complex molecules. Future developments will allow students to link structure of the molecules to physical and chemical properties.	Molecular visualization
Model-It	Modeling tool for dynamic systems; allows students to build, test, and evaluate dynamic qualitative models	Dynamic modeling
Web-It	HTML conversion tool to enable students to publish on the World Wide Web	Web publishing

The tools work together to support each phase of the inquiry process. When students are exploring the quality of a local stream, river, or lake, for instance, they can use probes (pH, temperature, dissolved oxygen, pressure) attached to portable technology and accompanying software to carry out collection of real-time data. The data can be uploaded to DataViz, where students can determine relationships and patterns using statistical analysis tools. By a variety of techniques such as digital photographs, graphs, and text, students visualize multiple types of data. In addition, student can link to representations and available animations to view the dynamic changes in different types of data.

Students can also use Artemis, an interface to the University of Michigan Digital Library. The digital library contains selected materials that are at appropriate levels of difficulty for middle and high school students. Supports in Artemis allow learners to sort, select, and organize documents and then easily return to them for further work.

Students can then use Model-It to build, test, and evaluate qualitative, dynamic models. For this, they import the functional relationships they developed in DataViz. They plan their models and create objects and factors. Using qualitative and quantitative representations, they then build relationship links

among the factors. A graphical view of each relationship is also provided. For visualization of the values of factors, Model-It provides meters and graphs. As students test their models they can change the values and immediately see the effects. Finally, students can use Web-It to publish their results on the Web.

Learner-Centered Design. We have learned a considerable amount about creating technology tools to support learners at different levels of expertise. The principles of learner-centered design (LCD) (Soloway, Guzdial, & Hay, 1994) recognize that students differ in a number of ways from professionals who use computational software. Students do not initially know the content they are exploring and must be supported as they engage in inquiry. They also differ from one another in technological expertise and in how they prefer to learn; the tools must therefore be adapted to different levels of complexity and represent information and data in multiple ways. And professionals in a field are more likely than students to be committed to their work. Technology must so design computational activities as to entice students to concentrate on substantive cognitive issues and problem solving.

The incorporation of learning supports or scaffolding that addresses the differences between learners and professionals is central to LCD. Scaffolding software enables the learner to achieve goals or accomplish processes that would otherwise be too difficult to obtain. Our software guides learners through steps within phases of inquiry; when constructing a model, for example, students are reminded to make a plan of variables to include before building and testing. Scaffolds support learners' metacognitive activities, prompting them to test individual relationships or a sequence of relationships before evaluating the entire model. The software supports testing and debugging, allowing students to determine which relationships work and which may need revision. Intrinsic scaffolding software supports different levels of expertise; it makes the simplest function available to the novice learner, but allows learners to gain access to advanced features as their capability grows. At first, for example, students build qualitative models; but as they gain experience they select a weighting tool to make their relationships quantitatively more precise.

## SUPPORTING PHASES OF INQUIRY

### Asking Questions

Good questions are both feasible to investigate and scientifically worthwhile, so that in exploring the answers students learn important science concepts. For

instance, one class was working on a project “Where does all our garbage go?” and conducting experiments on the effect of worms on decomposition. Groups of students were asked to generate sub-questions and to design their own investigations. One group of students asked the question: “Which types of material decompose and which don’t in light or dark, with worms?” This question allows students to explore important content addressed by the project and design an experiment to answer their question. To answer this question students needed to set up an experimental situation to test the impact of light on a decomposition environment. Several studies, however, have shown that initially student questions do not reflect these criteria (Erickson & Lehrer, 1998; King, 1990; Scardamalia & Bereiter, 1992).

One temptation among students, which Krajcik et al. (1998) found in seventh graders, is to choose their topics out of personal interest and preference. That is legitimate, even desirable, except that the interest may remain unrelated to the scientific merit of the questions. In the class working on the project “Where does all our garbage go?” one group of students asked: “When there is water in one [decomposition] bottle and apple juice in another which decomposed faster?” because one of the students liked apple juice. Only one student raised concerns about the merit of the question. It was not until the teacher conversed with the students about what apple juice might represent in nature that they realized that the experiment was about acid rain.

Students also fell back on personal experience. In a project on water quality one group asked the question: “Does the water in various places around Ann Arbor have fecal coliform? If so, then to what degree?” because one of the students in the group observed professionals testing his pond for fecal coliform. The investigation allowed students both to answer the question itself and to explore important content related to the project. Initially, however, the students did not seize on the scientific merit of the issue and the teachers needed to help them see it. The challenge we face as educators is how to capitalize on the personal experience of students as well as their interest, yet at the same time help the learners explore powerful scientific ideas.

There are several possible explanations for why students generate these types of questions. Students may not have enough experience with inquiry to fashion meaningful questions that are also feasible to carry out. In fact, students may view their task as generating a question that is acceptable to the teacher and capable of accomplishment in the classroom rather than as a task of building knowledge.

Deliberately allocating time and effort to identifying problems or questions, actively seeking explanations for them, and trying to build knowledge are behaviors that take time to develop in students. Bereiter (1990) has pointed up the difficulty. Some educators will claim that the way to promote the asking of good questions lies in traditional didactic instruction. But several studies show that asking good questions comes from experience at it and from learning how fruitful a well-thought question can be. Content and inquiry should be closely intertwined: a student's ability to generate questions is fostered through active engagement. In fact, Roth and Roychoudhury (1993) have reported that over time, as students engage in experimentation, the questions they use to guide inquiry become more specific and include particular variables and relationships. Scaradamaliala and Bereiter (1992) have found that the level of questions students pose improve as they explore a topic and gain more background knowledge. Krajcik et al. (1998) show that even early in their introduction to inquiry students are able to profit from suggestions of the teacher and also use what they have already learned.

Among strategies that have been shown to help students ask more productive questions is that proposed by Erikson and Lehrer (1998) of having the class as a whole develop critical standards for generating questions. Rather than imposing definitions of good questions, the teacher helps students to see that certain questions are less effective than others for building knowledge. For example, teachers might suggest that students should consider the breadth of the question, the ideas that need to be explored, the accuracy with which it reflects what students want to know, and the feasibility of finding information. Erikson and Lehrer demonstrate that over time students may develop critical standards that include interest along with potential for learning and for generating explanations and complex searches. The types of questions posed by the students the two investigated changed in accordance with these standards and came to require integration of multiple sources of information, selection of areas in which to search, and the generation of new questions. In attempting to answer these questions, moreover, students showed greater understanding and involvement.

To generate worthwhile questions, students must receive timely, informative, and critical feedback from teachers, peers, and others. They must also have opportunities to revise their questions and generate new ones. Whole class sessions can be committed to evaluating questions. To support such discussion and student self-assessment, teachers can provide skill templates as scaffolds. But in their concern to complete assigned work, students may be hesitant about devoting time to revising their work even when suggestions are offered. They may also

fail to understand how feedback can be used for improving their questions. Teachers must therefore emphasize the importance of revision and allocate time for it.

### Gathering Information: Inquiry on the World Wide Web

For seeking information to refine or answer a question, to design investigations, or to interpret findings, one increasingly popular source is the World Wide Web. Soloway, Krajcik, and their colleagues (Wallace et al., 1998; Hoffman & Wallace, 1997) have been exploring how students seek out and make use of information on the Web. Like others (Bereiter, 1990), they find that many students do not behave as intentional learners who aim to increase or build knowledge as they search. The task of seeking information about a question students often interpret as a matter of getting the right answer or good hits. Their background knowledge about their question, moreover, may be too limited to permit any keywords other than whatever is in their question. Failure to create synonyms may also be due to lack of appreciation for the significance of keywords or of understanding about how the technology works. Wallace et al. report that students do not have efficient ways of monitoring what they have accomplished; if the search continues over a period of time they lose their place, often repeating what they have done before or not making use of the information they have already gathered. Nor do they have sufficient strategies for reading or evaluating material online. Perhaps because students are used to looking up brief answers in textbooks or other reference sources, they may have neither skills nor inclination to criticize what they find (Mergendoller, 1996).

These findings point to some of the assistance students need if they are to conduct effective searches, and suggest that a major challenge to using digital information resources is to provide tools that enable the students to embed information seeking in a sustained process. Such tools must support both searching for simple facts and complex exploration of information when learners are trying to understand a multifaceted problem.

Classroom observations of students led to the creation of Artemis as an aid to students as they access and use digital information over the World Wide Web (Wallace et al., 1998). Artemis allows students to accomplish multiple tasks within a single computer environment. This keeps work from becoming fragmented and permits users to return to where they left off in prior sessions. The workspace provides for recording of searches and includes links to actual documents, helping students sustain the search process over time.

One feature, the question folders, supports students in thinking about and organizing the information they find in a way that most effectively addresses their query. They also help students to note what other questions or information they might usefully pursue. Students can store in question folders links to items they find interesting and can create multiple folders that reflect different components of the search or the refinements of an initial question. The folders allow flexibility in storing links and are available across many work sessions so that students can draw on what they have done before. Students can add or delete items or evaluate what they have found to date. Windows of results keep a live list of student searches so that they can see how they searched previously and what they have found. Observations indicate that students forget which queries they have submitted, and consequently repeat the same questions.

A broad feature includes a list of topics organized by domain. The topics present a hierarchy of terms that can be browsed or searched as the first step in creating a query. The feature is intended to help students generate keywords and draw upon prior knowledge as well as giving them a view of the structure of the content area they are exploring and providing them with alternative and productive ways to search.

Artemis is connected to the University of Michigan Digital Library, which contains a collection of relevant sites for middle grade students (see <http://umdl.soe.umich.edu>). The objective is to alleviate the frustrating problem students often have of getting numerous irrelevant hits in a Web search. Teachers and students also have the ability to criticize and recommend sites. Reading others' recommendations and their accompanying rationales, and contributing their own critiques can help students learn to evaluate information and sites, besides increasing motivation.

### Designing and Planning Investigations

Krajcik et al. (1998) found that during their initial experience with designing investigations, middle school students created experimental and descriptive designs differing in complexity from using only one variable to comparing several levels of different variables. Small group discussions about designs primarily centered on feasibility and procedures. Many students, for instance, considered the types of samples to use, ways to create or obtain the samples, and the amount of material needed. They also discussed the need for controls. Some groups, however, had difficulty grasping how to create controlled environments, confounded variables, and misjudged the feasibility of what they were trying to

do. The students' planning for data collection ranged from thoughtful to haphazard. Good plans included measurements related to the question, specifying what students were looking for as they measured or observed, and indicating the number of times measurements would be taken. They also detailed procedures to follow and included a way that data would be tracked and organized. Some planning problems students had were with qualitative techniques that involved drawing or writing a brief statement of what they observed. Generally, groups specified neither what they were looking for nor how the observations would help in answering their questions. Those who planned to use quantitative data often included measures with which they were familiar, like pH, but were not always appropriate for their purposes.

Students who have had little experience at gathering or interpreting data are not proficient at eliminating uninformative measures and do not realize the importance of being clear about their purpose. Thus students would benefit from having to explain how the measures selected relate to their questions, and be specific about what particular observations will indicate about the problem under study. Students need help in creating realistic plans. They sometimes overestimate how much they can accomplish within the time allocated and run out of time to complete the complicated set of things they have decided to do.

Although the students observed by Krajcik et al. presented their designs and plans to the class for suggestions, the presentations as well as many of the comments concerned specifics of the procedures rather than their purpose. Templates, which include questions about how the design and measures answer the question, could be used to guide the content of presentations and of the questioning that might accompany them. Such templates will also make it possible for peers to plan or students to engage in self-assessment of their plans. Allowing time for students to incorporate feedback, revise their plans, and emphasize the scientific merit of the inquiry is crucial in helping students create better design and plans.

### Carrying Out Investigations

It is important for students to be thorough, systematic, and precise in collecting and describing data. Krajcik et al. (1998) report that many students were careful in setting up experimental procedures and constructing apparatus, following directions precisely. But though many were quite careful to create charts to help them track and organize data, they varied considerably in how systematic they were in following through on their plans. Some groups ran out of time because they did not share responsibility for data collection and consequently failed to

complete necessary measurements. Others did not collect the measures they had planned but fixed on phenomena that attracted their attention, like bad smells or strange looking molds. They did not indicate how these phenomena were related to the scientific issues under study.

These problems illustrate students' need for help in managing complexity and time, and centering their attention on both the inquiry question and on the immediate needs entailed in collecting data. Students especially had trouble when plans called for numerous observations or complex procedures for data collection. Often students did not specify what they should record when collecting qualitative data or the reason for recording it. Doing both probably would help students focus on what data are important to gather. When first introduced to inquiry, moreover, students may not appreciate the need for consistency in measurement, following through on procedures, or maintaining experimental controls.

One solution for helping students handle complexity is to simplify and specify procedures so that learners can think about content. But even in more structured laboratory experiments, students tend to concentrate more on coping with procedures than on what they are supposed to learn (Hofstein & Lunetta, 1982). Students also need opportunities to think about how to make procedures precise and complete. Perhaps most challenging is that when students generate questions that are multifaceted, educators need to determine ways to help them reduce the complexity of the phenomena under investigation so that they can manage the work without compromising the integrity of the science and the authenticity of the problem.

Student interest in incidental observations must be for teachers an occasion for explaining how they bear on the larger scientific concepts under study. Krajcik et al. (1998) report that students were excited about what they were building and frequently asked about one another's work. They occasionally had animated conversations about unexpected changes that attracted their attention and tried to find out more about what they saw. But they rarely pursued the scientific implications of the observations or considered what they might suggest about other related questions or investigations. Surprise and curiosity can be an initial step in heightening interest in the work (Renninger, Hidi, & Krapp, 1992). The teacher should be ready to turn such moments into sustained cognitive engagement.

Microcomputer-based laboratories (MBL) can reduce complexity of data collection and representation that interfere with students' thinking about conceptual aspects of the inquiry. Students can use probes to monitor the temperature of a pond, to measure the pH of the pond, or to determine how dissolved

oxygen varies at different locations in it. Although many of these measurements can be done with traditional laboratory equipment, using MBL has a number of advantages. Probes can save time. They are also more reliable instruments. They can display the results both graphically and numerically so that children can more easily interpret the findings. A major advantage is the simultaneous collection and graphing of data visually and numerically, which contributes to the students' understanding (Brasell, 1987; Mokros & Tinker, 1987). Another advantage is that probes allow students to do explorations not typically possible in the science classroom. For instance, using a temperature probe, students can continuously track the temperature of a decomposition column. They can answer questions like, "Does the temperature of the column change at night?" or set up an experiment in an aquarium to see whether dissolved oxygen changes with amount of light.

### Analyzing and Interpreting Data

Krajcik et al. (1998) note that though the students had prepared charts and tables to record and organize their observations, they did not make graphs of quantitative data or create summary columns of qualitative data to facilitate comparisons across time and conditions even when teachers suggested that they do so. Perhaps because students did not look for patterns, their reports provided little interpretation. Instead, they tended to list findings with minimal elaboration, and failed to articulate how they had arrived at conclusions or to create logical arguments in which data were used to justify conclusions. Nor did they consistently draw upon background information to help interpret their findings. Penner et al. (1998) also tell of students who in creating models tended to describe data rather than identify principles that had produced them. Linn (1992) notes similar omissions; students using Computers as Learners Partners experienced difficulties using the results of laboratory experiments to explain everyday experience, and relied instead on intuitive ideas rather than the ideas under study.

One reason for this problem may be that students have had limited experience with these tasks and also may not know how to develop logical arguments to support their claims. Coleman (1998) reports on students who judged explanations as scientific if they included information that not everyone knew or could see with their own eyes, or information that needed to be discovered rather than looked up in a book. Palinscar, Anderson, and David (1993) have shown that students need considerable assistance in the process of argumentation, and have developed a program to help them systematically consider

alternative explanations for phenomena and to provide justification for their reasoning.

Teachers, it is clear need to model how students might go about the process of data analysis and interpretation. But many teachers may not have experience with this phase of inquiry; they are more likely to have dealt with data from highly structured laboratory experiments where the findings are known ahead of time. Exciting new software tools are now available to support students in interpreting data. Model-It allows them to build models that illustrate qualitative and quantitative relationships among data. In developing models, students specify objects and articulate relationships. As they construct and revise models students examine patterns and trends in data and consider the match between the phenomenon under study and the model they have created. DataViz enables students to link various data types together; for instance, pictorial data can be connected to numeric data. Viewing data in these new ways may help enhance student understanding. Evidence from several studies indicates that students can build fairly complicated and accurate models that illustrate deep understanding of science concepts and their relationships (Stratford, Krajcik, & Soloway, 1998; Spitulnik et al., 1997).

## ROLE OF METACOGNITION IN INQUIRY

Metacognition or self-regulation involves planning a course of action, monitoring progress to determine whether goals are being reached efficiently and effectively, and evaluating whether a change in plans or approaches is warranted. To stay organized students must track progress and stay focused on their problem, rather than getting confused or sidetracked by its elements. Doing so requires tactical and strategic metacognition. The tactical need is for regulation of cognition so that students can monitor their thinking as they work through details of tasks, such as who will be responsible for collecting data or using all the data collected in creating models. More strategically, students must think through what might seem to be disconnected elements to organize their efforts in service of the large purpose of the inquiry, such as how the data collection relates to the driving question, what data might be omitted if time runs out, or in what ways the model generated represents an answer to the driving question rather than just a representation of the data. Both types of thinking are needed for students to be systematic, accurate, and thorough and to make appropriate modifications or to adjust their strategies during inquiry; otherwise investigation runs the danger of becoming more like activity-based science where connections among activities and links to the overall issue or question often are not evident.

White and Frederiksen (1995) have explored ways to promote metacognition during inquiry. (See also their chapter in this collection.) They argue that metacognitive competence requires students to acquire the language to recognize and report on cognitive activities. ThinkerTools Curriculum provides this language through seeding conversations with categories chosen to represent metacognitive functions such as reflection on goals and on process. Examples of language for goals include formulating hypotheses and designing investigations. Labels are designed to help students recognize, monitor, and communicate about cognitive activities like generating multiple options or employing systematic strategies. Each of these can be further broken down into particular strategies and methods that are employed in each stage of the research process. For instance, being systematic means being careful, organized, and logical in planning and conducting work. When problems come up, helping students focus on their thought processes promotes effective decision-making. Students also use these criteria to do reflective assessments in which they evaluate their own and their classmates' research. White and Frederiksen (1995) have shown that engaging in reflective assessment enhances students' understanding of content and of science inquiry, and is especially beneficial for low achievers.

## THE ROLE OF COLLABORATION IN INQUIRY

The aim of collaboration is to build communal knowledge through conversation. It can occur within a whole class, among groups in a class, and with people and groups outside the classroom. Collaboration helps students construct knowledge and introduces them to disciplinary language, values, and ways of knowing. As students converse, they must articulate their ideas clearly, and consider and draw on the expertise of others (Bruer, 1994). In collaborations of this sort, groups are not as highly structured as are small cooperative groups. The aim is to share ideas with the whole class or community in order to enhance knowledge of all individuals. In contrast to cooperative learning programs, there is little emphasis on assigning roles, group rewards, or group competitions (see Slavin, 1990; Blumenfeld et al., 1996).

Effective collaboration requires students to share ideas, take risks, disagree with others and listen to them, and generate and reconcile points of view. As they work together, they must manage substantive, procedural, and affective matters. Often they direct attention to the latter two concerns rather than to substantive issues (Anderson, Holland, & Palincsar, 1997). Students do not spontaneously or naturally generate highly efficient questions or explanations on

their own and do not productively evaluate or respond to the explanations of others. Attempts to promote interactions include instructing in listening, resolving conflicts, and appreciating the skills and abilities of others (Webb & Palincsar, 1996).

One popular approach to facilitating student discussion and comprehension is reciprocal teaching (Palincsar & Brown, 1984; Rosenshine, Meister, & Chapman, 1996). As students read, teachers raise aloud a series of questions such as “What is likely to happen next? What do we know already?” Teachers model how more expert readers deal with text, eventually releasing responsibility to learners. Coleman (1998) has used conversational aids to improve the discussion of ideas and quality of explanations. During small group sessions, students used these prompts as they constructed concept maps. For instance the prompt, “Can you explain this in your own words?”, encouraged students to construct explanations. Another prompt, “Can you explain why you think this answer is correct?”, brought them to justify their responses. “Can you explain this using scientific information learned in class?” induced students to draw on background knowledge. Although they clearly benefited from such supports, the prompts did not always engender productive discussion; at times no one responded to the prompt, or the discussion went off track, or the discussion did not result in an explanation. Even when such conversational aids are employed, teachers need to monitor groups carefully.

Several tools are also available to promote collaboration and improve the quality of discourse (Pea & Gomez, 1992; Songer, 1998). The Computer Supported Intentional Learning Environment (CSILE), developed by Scardamalia and Bereiter (1991), advances understanding of subjects through electronic conversations centered on building a common database. CSILE has been used to support student investigations of topics such as endangered species, fossil fuels, evolution, and human biology. At the beginning of the year CSILE is empty; throughout the year it is populated by students’ contributions of text and graphical notes. The electronic database includes four categories of notes or thinking types. These categories correspond to stages in the investigation process. The first two, “what I know” and “high-level questions,” are used at the beginning of an investigation. They then use “plans” to generate a strategy for proceeding, and “new learning” to build a knowledge base. Their notes as they proceed in gathering information can be in text or graphical form. They can be commented on or added to by other students. The notes are structured to aid student conversation. They include opening phrases like, “One thing I don’t understand is...” or “A reference I thought you might find

useful is. . . .” The purpose is to assist students in asking further questions, raising counter arguments, suggesting additional sources of information, or offering feedback. Ultimately students write reports that synthesize the results of the class’s investigation.

## CONCLUSION

Achieving understanding of science concepts and processes requires supports through each phase of inquiry. These can come from a variety of sources—from the teacher, from curriculum materials, from technology, and from peers within and outside the classroom.

In almost all cases, the supports described here are designed to encourage students to be thoughtful as they explore ideas through investigation. Inquiry demands tactical self-regulation with respect to particulars, such as generating a question or designing an investigation, examining whether the question will actually allow for exploration of the problem at hand or whether the design is adequate for generating useful information. It demands as well long-term self-discipline in setting goals and making modifications according to constraints like time and resources, or to discoveries that might result in revision of designs or procedures. Initially, students will lack experience at being such intentional learners and are likely to need a great deal of assistance.

Inquiry also poses challenges for teachers. It requires different types of instruction that give equal weight to promoting thinking and teaching content, different management routines as groups of students work on various aspects of phases of inquiry, different ways of promoting student interaction and conversation, and different ways of monitoring student progress and understanding. It also necessitates different uses of time—for scaffolding, feedback, discussion and sharing, revision, and reflection. Allocating time can be disconcerting for teachers who worry about curriculum coverage. Nevertheless, although teachers at first find difficult a pedagogy based on inquiry, they report considerable satisfaction in seeing students motivated to learn, becoming proficient at asking questions and devising ways of answering them, and demonstrating deep understanding of scientific concepts. Sharing experiences and continued exploration of techniques is essential to meeting the recommendation of the National Research Council (1996) that inquiry become a predominant mode of instruction.

## ENDNOTES

1. The authors would like to thank Ann Rivet from the University of Michigan for her helpful editorial comments.
2. These principles are being instantiated into the curriculum by a team including Margaret Roy, Jon Singer, and Becky Schneider from the University of Michigan and Karen Amati from the Detroit Public Schools.
3. The development of the Investigators' Workshop has been supported by grants from the National Science Foundation (REC 9554205 and 955719).

## REFERENCES

- American Association for the Advancement of Science. 1993. *Benchmarks for science literacy*. New York: Oxford University Press.
- Anderson, C.W., J.D. Holland, and A.S. Palincsar. 1997. Canonical and sociocultural approaches to research and reform in science education: The story of Juan and his group. *Elementary School Journal* 97: 360-383.
- Atwater, M.M. 1994. Research on cultural diversity in the classroom. In *Handbook of research on science teaching and learning*, edited by D.L. Gabel, 558-576. New York: MacMillan.
- Audet, R.H., P. Hickman, and G. Dobrynina. 1996. Learning logs: Classroom practice enhancing student sense making. *Journal of Research in Science Teaching* 33: 205-222.
- Bass, K.J., and G.P. Baxter. 1998. Writing in science...for what purpose? An analysis of notebooks in hands-on science classrooms. Paper presented at the annual meeting of the American Education Research Association, San Diego.
- Bereiter, C. 1990. Aspects of an educational learning theory. *Review of Educational Research* 60: 603-624.

- Blumenfeld, P.C., E. Soloway, R. Mark, J. Krajcik, M. Guzdial, and A. Palincsar. 1991. Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist* 26: 369-398.
- Blumenfeld, P.C., R.W. Marx, J.S. Krajcik, and E. Soloway. 1996. Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher* 25(8): 37-40.
- Blumenfeld, P.C., R.W. Marx, H. Patrick, and J.S. Krajcik. 1998. Teaching for understanding. In *International handbook of teachers and teaching*, Vol. 2, edited by B.J. Biddle, T.L. Good, and I.F. Goodson, 819-878. Dordrecht, The Netherlands: Kluwer.
- Brown, A.L., and J.C. Campione. 1994. Guided discovery in a community of learners. In *Classroom lessons: Integrating cognitive theory and classroom practice*, edited by K. McGilly, 229-270. Cambridge, MA: MIT Press/Bradford Books.
- Britsch, S.J., and D.P. Shepardson. 1997. Children's science journals: Tools for teaching, learning and assessing. *Science and Children* 34: 12-17.
- Brasell, H. 1987. The effect of real time laboratory graphing on learning graphic representation of distance and velocity. *Journal of Research and Science Teaching* 24(4): 385-395.
- Bruer, J. 1994. Classroom problems, school culture, and cognitive research. In *Classroom lessons: Integrating cognitive theory and classroom practice*, edited by K. McGilly, 273-290. Cambridge, MA: MIT Press.
- Cognition and Technology Group at Vanderbilt. 1992. The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist* 27: 291-315.
- Coleman, E.B. 1998. Using explanatory knowledge during collaborative problem solving in science. *The Journal of the Learning Sciences* 7 (3&4): 387-428.

- Collins, A., J.S. Brown, and S.G. Newman. 1989. Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, edited by L.B. Resnick, 453-494. Hillsdale, NJ: Erlbaum.
- Edelson, D. 1998. Realizing authentic science learning through the adaptation of scientific practice. In *International handbook of science education*, edited by B.J. Fraser and K.G. Tobin, 317-331. Dordrecht, The Netherlands: Kluwer.
- Erickson, J., and R. Lehrer. 1998. The evolution of critical standards as students design hypermedia documents. *The Journal of Learning Science* 7 (3&4): 351-389.
- Fensham, P., R. Gunstone, and R. White. 1994. *The content of science: A constructivist approach to its teaching and learning*. London: Falmer Press.
- Hmelo, C.E., and S.M. Williams. 1998. Learning through problem solving. Special Issue. *The Journal of Learning Science* 7(3&4).
- Hoffman, J., and R. Wallace. 1997. Structuring on-line curriculum for the science classroom: Design, methodology, and implementation. Paper presented at symposium, Using online digital resources to support sustained inquiry learning in K-12 science. Annual meeting of the American Educational Research Association, Chicago.
- Hofstein, A., and V.N. Lunetta. 1982. The role of the laboratory in science teaching: Neglected aspects of research. *Review of Educational Research* 52: 201-217.
- Hyerle, D. 1996. *Visual tools for constructive knowledge*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Jackson, S., S. Stratford, J. Krajcik, and E. Soloway. 1996. Making system dynamics modeling accessible to pre-college science students. *Interactive Learning Environments* 4: 233-257.

- Jackson, S., J. Krajcik, and E. Soloway. 1999. Model-It: A design retrospective. In *Advanced designs for the technologies of learning: Innovations in science and mathematics education*, edited by M. Jacobson and R. Kozma, pgs. Hillsdale, NJ: Erlbaum.
- Keys, C.W. 1994. The development of scientific reasoning skills in conjunction with collaborative writing assignments: An interpretive study of six ninth-grade students. *Journal of Research in Science Teaching* 31: 1003-1022.
- King, A. 1990. Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal* 27: 664-687.
- Krajcik, J.S., C.M. Czerniak, and C. Berger. 1999. *Teaching children science: A project-based approach*. Boston: McGraw-Hill.
- Krajcik, J., P. Blumenfeld, R.W. Marx, K.M. Bass, J. Fredericks, and E. Soloway. 1998. Middle school students initial attempts at inquiry in project-based science classrooms. *The Journal of Learning Sciences* 7 (3&4): 313-350.
- Krajcik, J.S., P.C. Blumenfeld, R.W. Marx, and E. Soloway. 1994. A collaborative model for helping teachers learning project-based instruction. *Elementary School Journal* 94: 483-497.
- Ladson-Billings, G. 1995. Toward a theory of culturally relevant pedagogy. *American Educational Research Journal* 32: 483-491.
- Linn, M. 1992. The computer as learning partner: Can computer tools teach science? In *This year in school science 1991: Technology for teaching and learning*, edited by K. Sheingold, L.G. Roberts, and S.M. Malcom, 31-69. Washington, DC: American Association for the Advancement of Science.
- Linn, M.C. 1998. The impact of technology on science instruction: Historical trends and current opportunities. In *International handbook of science education*, edited by K. Tobin and B.J. Fraser, 265-294. Dordrecht, The Netherlands: Kluwer.

- Lunetta, V.N. 1998. The school science laboratory: Historical perspectives and contexts for contemporary teaching. In *International handbook of science education*, edited by K. Tobin and B.J. Fraser, 249-264. Dordrecht, The Netherlands: Kluwer.
- Marx, R.W., P.C. Blumenfeld, J.S. Krajcik, and E. Soloway. 1998. New technologies for teacher professional development. *Teaching and Teacher Education* 14: 33-52.
- McGilly, L. 1994. *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: MIT Press.
- Mergendoller, J.R. 1996. Moving from technological possibility to richer student learning: Revitalized infrastructure and reconstructed pedagogy. *Educational Researcher* 25: 43-46.
- Metz, K.E. 1995. Reassessment of developmental constraints on children's science instruction. *Review of Educational Research* 65: 93-128.
- Mokros, J.R., and R.F. Tinker. 1987. The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching* 24: 369-383.
- National Research Council. 1996. *National science education standards*. Washington, DC: National Academy Press.
- Newmann, F.M., and G.G. Wehlage. 1993. Standards for authentic instruction. *Educational Leadership* 50:8-12.
- Newmann, F.M., and D.A. Archibald. 1992. Approaches to assessing academic achievement. In *Toward a new science of educational testing and assessment*, edited by H. Berlak, F.M. Newmann, E. Adams, D.A. Archibald, T. Burgess, J. Raven, and T.A. Romberg, 71-83. Albany, NY: SUNY Press.
- Novak, J.D., and D.B. Gowin. 1984. *Learning how to learn*. Cambridge, England: Cambridge University Press.

- Palincsar, A., and A. Brown. 1984. Reciprocal teaching of comprehension fostering and comprehension monitoring activities. *Cognition and Instruction* 1: 117-175.
- Palincsar, A.S., C. Anderson, and Y.M. David. 1993. Pursuing scientific literacy in the middle grades through collaborative problem solving. *The Elementary School Journal* 93: 643-685.
- Parks, S., and H. Black. 1992. *Organizing thinking: Graphic organizers*. Pacific Grove, CA: Critical Thinking Press and Software.
- Pea, R., and L. Gomez. 1992. Distributed multimedia learning environments: Why and how? *Interactive Learning Environments* 2 (2): 73-109.
- Penner, D., R. Lehrer, and L. Schauble. 1998. From physical models to biomechanics: A design-based modeling approach. *The Journal of Learning Sciences* 7 (3&4): 429-450.
- Perkins, D. 1992. *Smart schools: From training memories to educating minds*. New York: Free Press.
- Renninger, K.A., S. Hidi, and A. Krapp. 1992. *The role of interest in learning and developments*. Hillsdale, NJ: Erlbaum.
- Rosenshine, B., C. Meister, and S. Chapman. 1996. Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research* 66: 181-221.
- Roth, W.M. 1995. *Authentic school science*. Dordrecht, The Netherlands: Kluwer.
- Roth, W.M., and A. Roychoudhury. 1993. The development of science process skills in authentic contexts. *Journal of Research in Science Teaching* 30: 127-152.
- Salomon, G., D.M. Perkins, T. Globerson. 1991. Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher* 20: 2-9.

- Santa, C.M., and L.T. Havens. 1991. Learning through writing. In *Science learning, processes and applications*, edited by C.M. Santa and D.E. Alvermann, 122-133. Newark, DE: International Reading Association.
- Scardamalia, M., and C. Bereiter. 1991. Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences* 1: 37-68.
- Scardamalia, M., and C. Bereiter. 1992. Text-based and knowledge-based questioning by children. *Cognition and Instruction* 9: 177-199.
- Sherwood, R., A. Petrosino, X.D. Lin, and the Cognition and Technology Group at Vanderbilt. 1998. Problem based macro contexts in science instruction: Design issues and applications. In *International handbook of science education*, edited by B.J. Fraser and K. Tobin, 349-362. Dordrecht, The Netherlands: Kluwer.
- Singer, J., J. Krajcik, and R. Mark. 1998. Development of extended inquiry projects: A collaborative partnership with practitioners. Paper presented at the annual meeting of the American Educational Research Association, San Diego.
- Slavin, R.E. 1990. *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice Hall.
- Soloway, E. 1997. Using on-line digital resources to support sustained inquiry learning in K-12 science. Symposium presented at the annual meeting of the American Educational Research Association, Chicago.
- Soloway, E., M. Guzdial, and K.E. Hay. 1994. Learner-centered design: The challenge for human computer interaction in the 21st century. *Interactions* 1: 36-48.
- Soloway, E., and J.S. Krajcik. 1996. *The investigator's workshop: Supporting authentic science inquiry activities*. Washington, DC: National Science Foundation.

- Songer, N.B. 1993. Learning science with a child-focused resource: A case study of kids as global scientists. In *Proceedings of the Fifteenth Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
- Songer, N.B. 1998. Can technology bring students closer to science? In *International handbook of science education*, edited by B.J. Fraser and K.G. Tobin, 333-347. Dordrecht, The Netherlands: Kluwer.
- Spitulnik, M.W., S. Stratford, J. Krajcik, and E. Soloway. 1998. Using technology to support student's artifact construction in science. In *International handbook of science education*, edited by B.J. Fraser and K.G. Tobin, 363-381. Dordrecht, The Netherlands: Kluwer.
- Stratford, S.J., J. Krajcik, and E. Soloway. 1998. Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology* 7(3): 215-234.
- Wallace, R., E. Soloway, J. Krajcik, N. Bos, J. Hoffman, H.E. Hunter, D. Kiskis, E. Klann, G. Peters, D. Richardson, and O. Ronen. 1998. ARTEMIS: learner-centered design of an information seeking environment for K-12 education. In *Human Factors in Computing Systems, CHI-98 Conference Proceedings*, 195-202.
- Webb, N.M., and A.S. Palincsar. 1996. Group processes in the classroom. In *Handbook of research in educational psychology*, edited by D. Berliner and R. Calfee, 841-873. New York: Simon & Schuster.
- White, B.Y., and J.R. Frederiksen. 1995. *The ThinkerTools inquiry project: Making scientific inquiry accessible to students and teachers*. Berkeley: University of California, School of Education. Causal Models Research Group Report No. 95-02.
- White, R., and R. Gunstone. 1992. *Probing understanding*. London: Falmer Press.

## Constructing Scientific Models in Middle School

Karen Amati

---

Located in Detroit, Michigan, Lessenger Middle School is a Title 1 school with about 850 students, most of them African American and from single parent homes. When Dr. Joseph Krajcik, a professor from the University of Michigan, suggested we pilot a scientific modeling program, my initial response was “No way!” Scientific models are built upon statistics, and our students have a difficult time with fundamental word problems. How could children model their understanding of science? How could constructing a model help them better understand science? Isn’t it necessary to understand thoroughly a topic before constructing a model? When we examined models constructed by children, the program spoke for itself. Because the models used verbal quantifiers such as “some,” “a lot” and “none,” I realized that our children could do these.

Our first scientific use of Model-It was as a culminating assessment of the understanding of weather our students had gained in Kids as Global Scientists. Dr. Nancy Songer, the project investigator, and the co-director Dr. Perry Samson along with their staff analyzed our students’ pretests and posttests and found that they excelled over participants from schools across the country. The curriculum materials were essentially the same for all of the groups. The difference was the use of Model-It. This was impressive! We are not accustomed

to seeing our students excel over children from school systems that experience much higher test scores than we.

Students and teachers view Model-It as a tool to make sense of the many pieces of information acquired in the process of exploring a topic. Although concept maps also help students make meaning, Model-It takes it one step further allowing students to test their ideas. The *National Science Education Standards* (National Research Council, 1996) stress that students need to make meaning rather than just memorize information. As the science and technology resource teacher, I have helped teachers successfully use Model-It in grades six through eight, and I was thrilled to find that students remembered how to use the program the following year.

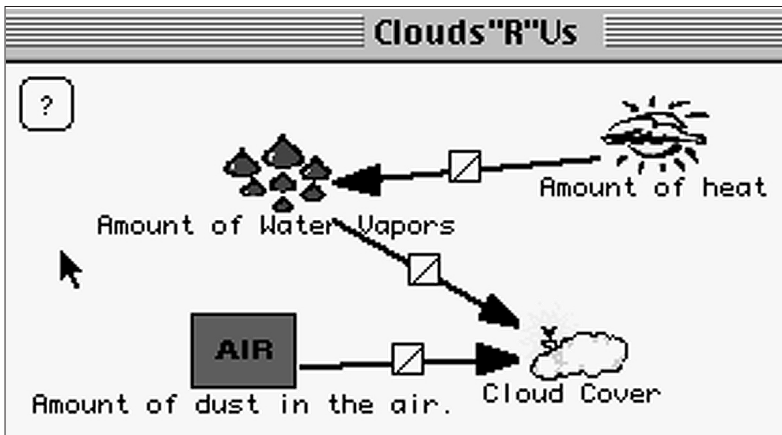
## WHAT IS MODEL-IT?

Model-It is a computer program designed and developed by Shari Jackson when she was a graduate student in Computer Science at the University of Michigan and by two professors from the University of Michigan, Elliot Soloway, in Educational Technology and Joseph Krajcik in Science Education. The program allows students to construct computer models of complex, dynamic science systems using graphics and verbal descriptions. Mathematical quantification is possible. At first glance, the scheme seems to be an elaborate concept map. Concept maps, however, are static; students have no way to verify relationships. Model-It is dynamic. Students can easily test, analyze, and revise their models, and that helps students make meaning of the phenomena they are modeling. Another advantage of the program is that detailed descriptions must be written for each relationship. The descriptions are much more effective than the word or two used in a concept map. In the paragraphs here, each section of Model-It will be described and illustrated by screen shots of student models.

## WHAT DOES A MODEL LOOK LIKE?

This is the beginning of a model constructed by a sixth-grade group called “Clouds ‘R Us.” Each group of students became topic specialists. Prior to beginning the weather unit, they studied the phases of matter and the effect of heat on evaporation. During the first part of the weather unit, they became experts on clouds. Following the arrows, this model shows that as the amount of heat from the sun increases, the amount of water vapor increases. Another arrow indicates that as the amount of water vapor increases, the amount of

cloud cover increases. The bottom arrow shows that as the amount of dust in the air increases, so too does the amount of cloud cover. The model demonstrates that students were able to use various concepts to construct complex relationships of cause and effect concerning one facet of weather. The model graphic does not reveal the depth of understanding required to construct a model. Some students, for example, may consider the model to be complete without the dust factor. The teacher may ask them whether they think clouds form whenever there is water vapor. Through continued questioning the students will recall that dust particles are needed. When they define the relationship, they cannot simply state that as dust particles increase the amount of cloud cover increases. They must explain that the dust particles act as a nucleus for gathering moisture to form a droplet of water.



## WHAT IS THE PROCESS OF CREATING A MODEL?

### Defining Objects and Factors

First the students, either individually or in groups, draw a picture of those things that affect weather. Next, the teacher helps them consolidate into a manageable number of objects the things in their drawings. In the model above, the objects are air, weather, sun, and water. After the objects are identified, the students choose an object and a measurable factor. These students chose the sun as the object and the “Amount of heat” as a factor of the effect of the sun on weather. Notice that text rather than numbers quantifies the factor. The students are free

to be creative when they discuss the science concepts and choose the best words for describing the maximum, medium, and lowest amount of the factor. The descriptors are entered in the three text range boxes. In this case, students chose A lot, Some and None. The dialog box calls for a description of the factor. We feel that class time should go to discussing the concepts rather than typing, which most of our students do only slowly. The descriptions are homework on an object/factor worksheet that must be completed before the students go to the computer. The two students working on a model compare their papers and choose the best parts of each one. Each student is required to enter factors into the model. The students will describe two or three factors and then move on to relationships or create many factors before linking them into relationships.

The image shows a software dialog box titled "Factor Editor". It has a standard Windows-style title bar with a mouse cursor pointing to the "Factor Editor" text. The dialog box contains the following elements:

- Object:** A dropdown menu with "Sun" selected.
- Factor:** A dropdown menu with "Amount of heat" selected.
- Initial Value:** A text box containing "Some".
- Text range:** Three vertically stacked text boxes containing "A lot", "Some", and "None". To the right of these boxes is a vertical scrollbar.
- Description:** A large empty text area.
- Buttons:** "Delete Factor", "Cancel", and "OK" are located at the bottom of the dialog.
- Icons:** A question mark icon and a sun icon are located in the upper right area of the dialog.
- Other:** An "Options..." button is located to the right of the "Initial Value" field.

### Building Relationships

The students return to their drawing of the things that affect weather. They are asked to look for pairs of objects that form relationships of cause and effect. This is easier than it might seem because it is a common task in our Language Arts classes. The students draw an arrow from the cause to the effect and explain why the relationship exists. Each student constructs some relationships on paper as homework. The computer partners discuss them and choose some from each paper. The dialog between the students, as they decide which are the best relationships, is a valuable opportunity for them to internalize the concepts. They must evaluate each others' information and offer evidence for their choices.

Anyone familiar with the operation of the program can click on the appropriate boxes and construct a relationship. The assessment criteria for an acceptable model require a detailed explanation for each. It is supposed to begin with a statement of the relationship followed by a "because" statement. This beginning model has only one acceptable explanation of a relationship.

The first relationship should explain why heat energy increases evaporation. This group would be encouraged to review its notes or one of its information sources to complete the explanation. A better relationship description would be that as the amount of heat increases the amount of evaporation increases because heat causes the liquid molecules to move faster and farther apart to become a gas.

The second relationship, that between water vapor and cloud formation, has a good explanation for sixth-grade students. The explanation would have been better if they had indicated that increasing moisture was responsible for the increasing amount of cloud cover.

The last explanation demonstrates a common error. Many of our students feel that a restatement of the relationship is an explanation or a justification for the relationship. Personal conferences with the students reveal that they do not understand the difference between a restatement of a question and an explanation. One of the greatest strengths of Model-It is that it forces students to complete a detailed explanation statement. Most of the students find this extremely difficult. Students who have succeeded in school because of their ability to memorize become especially frustrated. Conscientious students will agonize over the explanations, and pester their teachers with frequent requests for help with a relationship. One group nagged us for more than a week concerning the formation of hail. The students constantly quoted the memorized definition, but could not fit it into the relationship using the weather objects. They wanted a special object for hail. In our use of Model-It, students are required to state the relationship, add the word "because,"

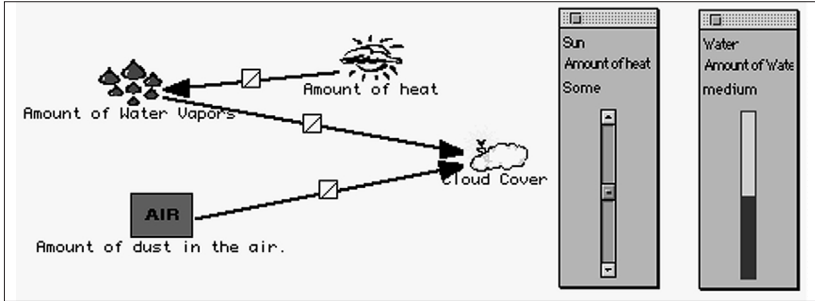
and end with an explanation. Usually, they cannot use memorized definitions or bits of information. For example, the last relationship could say, “As the amount of dust in the air increases, the amount of clouds increase because dust particles make it easier for water vapor to condense into water droplets to make clouds.” If students did not learn anything else during the unit, they learned that there is a world of difference between a memorized definition and understanding.

As teachers move from computer to computer, they observe only a sample of each group’s relationships. How many are incorrect? Are we allowing students to learn incorrect information? Many teachers express these concerns. Requiring as homework a paper version of the relationships allows the teacher to assess the quality of the contributions by each student and to target groups for increased individual attention. There are instructional advantages in that every student is personally engaged in relationship construction because students are not allowed to work at the computer until they do the paper version. Then in their workgroups they compare their relationships and either choose the best one or combine their ideas. This further compels them to think critically. The paper version is supportive of the teachers who are concerned that students will build incorrect relationships. Errors are valuable opportunities for reflection and learning. When groups having difficulties are targeted, teachers can have conferences with the students, and help them develop true relationships. For the teacher to mark the relationship wrong with the traditional red pen and lower the grade accordingly would be counterproductive. That is not the role of the facilitator. The evaluation of the content of the relationships should occur when the model is completed and the students have had adequate time to test their relationships.

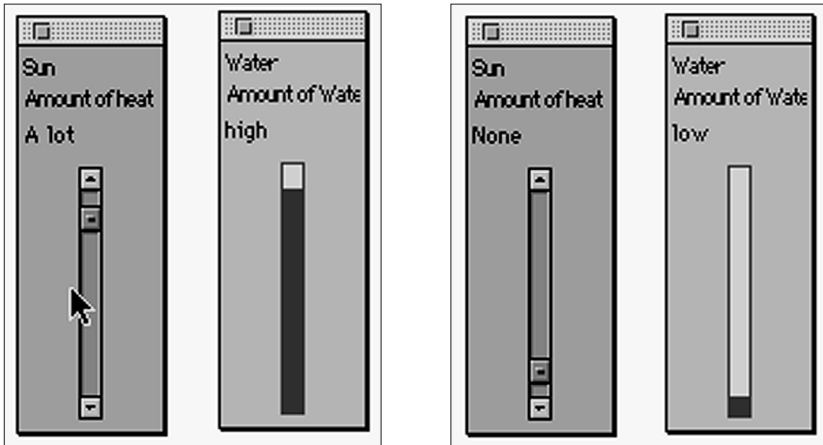
## Testing

The final phase of the modeling process is testing the relationships. Each relationship may be tested after it is constructed, or a number of relationships can be constructed and then tested.

In the example, the test meters show the initial midway settings using the quantifiers “Some” and “Medium” as chosen by the students. Note that the meters differ; one meter has a slider. The meter with the slider is the independent variable, the cause of change. The other meter is the dependent variable, the effect. Students are encouraged to test one relationship at a time to simplify troubleshooting should the relationship not work properly. Once the “Run” button is activated, students move the slider up and down changing the amount of the sun’s heat and noting the effect on the amount of water vapor.

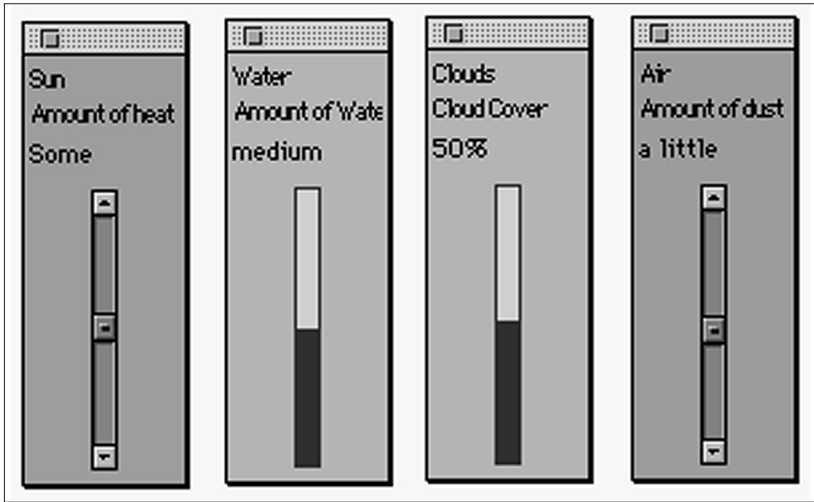


The slider first moved to the “A lot” position, causing the other meter to move to the “high” position. This indicates that if the sun’s heat is intense, the amount of water vapor will be high. The underlying assumption is that there is an ample supply of water.

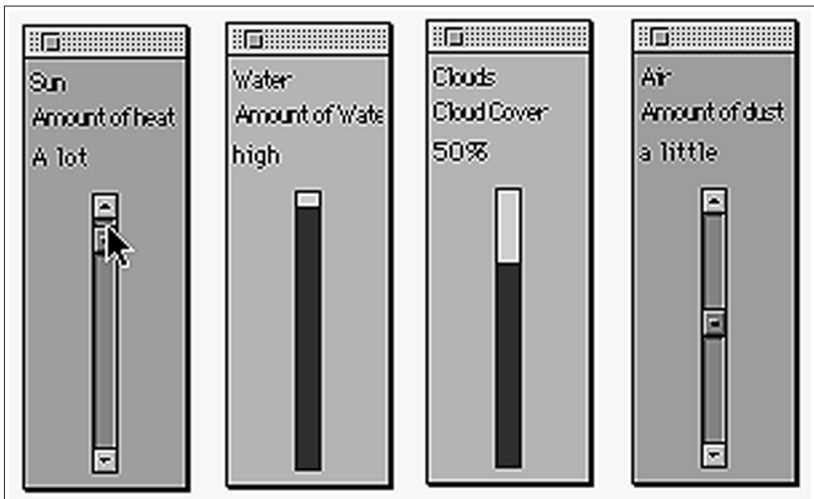


Next, the slider moved toward the bottom of the meter, causing the water vapor meter to decrease.

There are at least three different approaches to teaching the testing of relationships, and each requires its own level of sophistication. The simplest is to have the students open the meters, move them around and see whether the meters behave as they expected. A second level emphasizes the experimental nature of testing by having the students form hypotheses and test them with the meters. The most difficult approach identifies the independent and dependent variables. Middle school science standards expect students to be able to differentiate between variable types, and all but the brightest students find it very



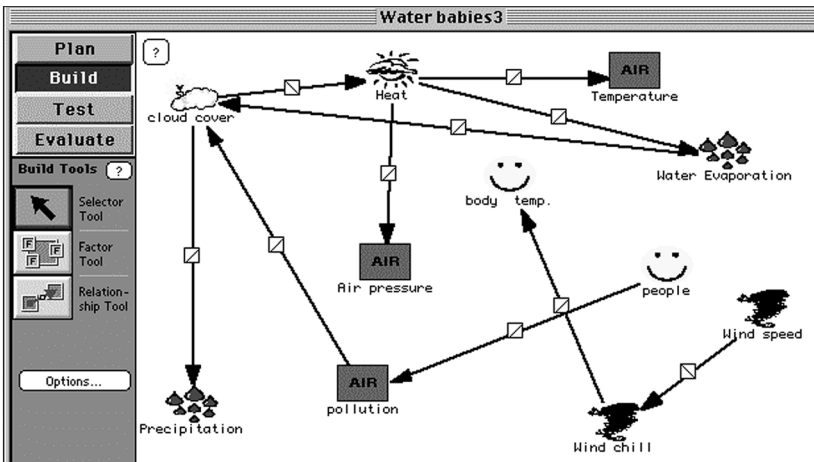
difficult. Model-It is an excellent tool for teaching the difference. The independent variable is the meter that can be changed freely, whereas the dependent variable does not have a slider and is moved by another meter. Repeated sessions of model building and testing allow multiple opportunities for students to practice identifying and explaining the difference between dependent and independent variables.



After pairs of relationships have been tested, the students increase the number of meters and may test all of them simultaneously. We begin with all of the meters set to their beginning levels. Next, the amount of heat is increased to the maximum, which causes the amount of water vapor to become high and the amount of cloud cover to increase. The students were not sophisticated enough to enter a numerical scale. The 50% is treated as a word rather than a number. Students read it as more than 50%. Many students expect the cloud cover to become 100% because the heat and water vapor are at their maximum values. Upon reconsideration and following a discussion led by a teacher, they realize that the amount of dust affects the cloud cover. In this very simple, beginning model, the amount of cloud cover does not move in a direct relationship to the amount of water vapor. The amount of dust particles also influences the amount of clouds. In the last set of meters, the dust particles have been increased and the cloud cover increased correspondingly. From this beginning level model, created by sixth-grade students, we can see that the construction and testing of models using Model-It as a tool allow students to discover the complexity of scientific systems.

### Group Presentations

The culminating activity is a group presentation to the rest of the class. Often presentations are viewed only as assessments. But students are more motivated to create an accurate, complex model when they know that they will be presenting it to their peers. Each student is required to participate actively in the



presentation by explaining at least one relationship. The screen shot is from a model presented by a group of seventh-grade students. The model and their presentation showed that they understood the water cycle, the very elusive concept of wind chill, and the difficult relationship between air pressure and temperature. Model-It gives students the freedom to express their knowledge using their own creativity and the movement of graphics. During the modeling activity, therefore, most students are actively engaged in learning science for nearly the whole class period each day. This model took several days to construct. Just imagine, seventh-grade students spending several hours willingly discussing such topics!

### MORE COMPLEX MODELS

These examples of examining weather lack one essential specification, the times at which the events of the model take place. In real conditions, not all of the events of the water cycle occur at one time. Model-It can accommodate the more complex question of whether the events happen immediately, quickly or slowly.

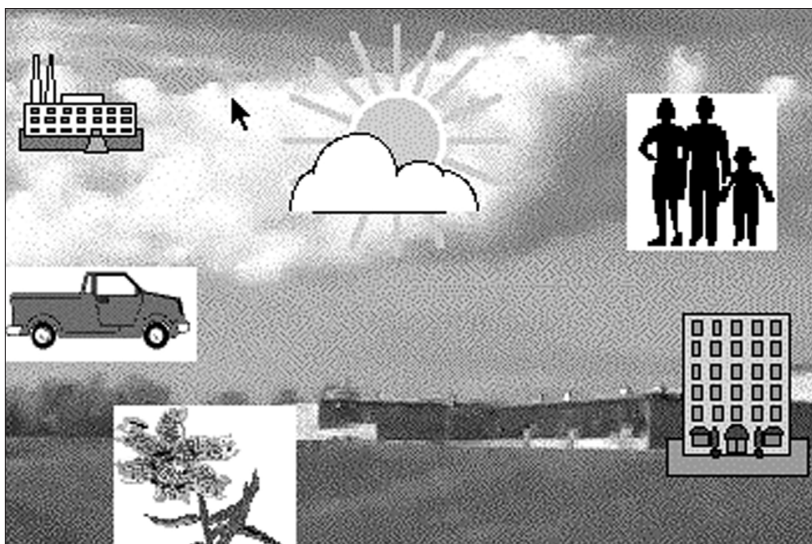
During our examination of the quality of our river water, the students use the weighting factor for each of nine water quality tests. The amount of dissolved oxygen in a river is many times more important than the turbidity. Model-It allows you to assign values that vary with the relationship between the two. This is indicated by the thickness of the relationship arrow. In the water quality model here, the students were given the nine weighted water quality relationships. We wanted them to concentrate not on the numerical values they worked with but on the factors in the environment of our river that affect each of the parameters: for example, the events that determine the amount of dissolved oxygen rather than the mathematical importance of dissolved oxygen. They were therefore given a template upon which they built their model. One of the problems was a lack of screen space for everything they wanted to include.

### MODELS DESIGNED BY TEACHERS

After using a number of model templates constructed by university personnel, I decided to construct one of my own. In the eighth grade, students review electrical circuits, learn about Ohm's Law, and study the relationships between electricity and magnetism. It was easy for me to draw a paper model that defined the objects needed for Model-It. Unable to find appropriate clip art, I had either to scan images or to take digital photos. I did the latter. A benefit to

the digital pictures is that the objects are readily recognizable by the students. Getting the artwork was the most difficult part. When the images were ready, they were pasted into Model-It in the Object Editor.

Many of our models have local pictures as one of the objects and as the background for the others. This helps to make projects tangible and relevant to the students. The air quality objects here are plants, vehicles, man-made objects, weather, people, and the background is a picture of the park and the back of our school. The background picture was taken with a digital camera. For other projects, I have asked the art students to draw the clip art I need. These were scanned and saved as clip art.



### CHOOSING THE QUESTION OR MODELING PURPOSE

Much of science does not fit into a complex system of cause and effect suitable for modeling. The question “What is the quality of the air in our community?” is the driving question for an authentic project-based science unit. Most of the objectives we wish to accomplish address differences between physical and chemical changes, word and formula equations, and the conservation of matter. The unit emphasizes the chemistry of air pollution. The environmental science aspects are secondary objectives and a means of making the study of chemistry applicable to the life and health of the students. However, the models address the environmental issues with an occasional very general reference

to a chemical aspect. Model-It did not foster deep understanding of the chemistry underlying the environmental science. “Do I really need to wear a helmet when riding my bike?” This is the driving question for another authentic project. Acceleration, velocity, inertia, and gravity have linear relationships to injury to a biker’s head, and only seldom will more than one cause point to an effect. Some of us argue that the construction of each relationship is a valuable learning experience because the students must convert their collection of knowledge into the graphic format. Others believe that there must be a network of relationships to make the testing phase meaningful and the modeling experience valuable. These arguments occur only because we are constantly mindful of the need in modeling projects to use every class minute to address efficiently the objectives of our curriculum.

Model-It is adaptable to any network of cause and effect. Social studies and history would be perfect subjects. Students could build models of the factors that brought about the Bill of Rights and the effects of the bill. Other topics that come to mind are slavery, westward expansion, and labor unions. One of our social studies teachers attended a series of Model-It training sessions, and I am looking forward to working with her to infuse Model-It into one of her units of study.

## STUDENTS AND TEACHERS ENJOY MODEL-IT

Students and teachers both enjoy their experiences with Model-It. Unlike many software applications, Model-It offers creativity to both. Students do not quickly tire of a process that actively involves them in the creation of a model representing them. While there are guidelines for the general content of the model, students are in control of the objects and factors chosen to represent relationships. Because they have control, they must make decisions. The decision making process requires them to use both vocabulary and reasoning. They have both the frustration and the support of working with a partner. They begin with a blank workspace that they must fill. Contrast this to a worksheet on which they fill in some empty spaces hoping to read the mind of its designer. Think about it. Why do we enjoy teaching? We enjoy the creative efforts of designing units of instruction and creative ways of presenting materials to our students. Model-It lets students experience the creative aspect of learning. They discover the teacher as a facilitator of learning, as another partner in the creation of their model, rather than as a taskmaster constantly making demands of them. The teacher therefore benefits too, changing from provider of information to facilitator of learning.

## TEACHER AS FACILITATOR

Most teachers are familiar with the emerging role of teachers as facilitators of learning. The classroom culture, however, prepares students, administrators, and parents to expect the teacher to tell the students what they need to know and to provide test results documenting their achievement. Model-It is an entirely new environment for all of the participants and an opportunity to break out of the old culture. It happens naturally. As students are constructing their models, the teacher is moving around the room solving technical problems or looking over shoulders and offering suggestions. The true facilitator does not tell the students what to do, but asks leading questions that bring the students to reconsider what they are doing. If students have a blatant error in their model, the teacher facilitator may ask them to explain the erroneous segment. As they do so, they often realize that there is an error. The teacher may ask questions to break the misconception into small segments, and further questioning will help the students assemble an appropriate component.

Once teachers have established this rapport with the class, the demand for their time becomes unmanageable. We have established the “Three Before Me!” rule. Students are to ask three people sitting near them before seeking the teacher’s help. This rule helps students to appreciate one another and raises the self-esteem of classmates who give assistance. Another technique that lowers everyone’s impatience for immediate help is the red and green paper cup signaling system. When the cup on the top of the computer is green, everything is going well. A red cup is a call for help. It does not take long for the students and teachers to come to appreciate this working atmosphere.

## COMPUTERS AS MOTIVATORS

Because the models are built on the computers, students are much more willing to meet the challenges of model building. It is commonplace to observe unmotivated or disruptive students actively participating in the modeling process. At these times, the discipline problems become minor or disappear. The computer also helps to address the differing kinds of intelligence found in any student group. The student who cannot comprehend the science may type well or have computer expertise and become valuable in the construction of the group’s model. Rejection of them for their limitations gives way to respect for their skill. In every class, there are fewer problems during the modeling phase than at any other time.

## PARENTS ARE FRUSTRATED

Parents find the modeling sessions beyond comprehension. Some are irate because they cannot understand factors and relationships. They want more written material so that they can work with their student. They want to know, "What page is it on?" The students readily adapt to changes involving computing, but computers intimidate many parents. We have printed screen shots of their models for them to take home and explain to their parents. Also, teachers invite parents to attend the science class and the final presentations. The best solution to this problem would be software for home, but the memory requirements are beyond those of lower priced computers.

## CONCLUSION

For more than fifteen years, I looked for software tools to enhance learning and teaching of science and mathematics. The computer is not a subject to be taught, but a tool to be used to facilitate a task. Over the years, I have had my favorite software titles, but none of them directly facilitated learning scientific principles. They were excellent tools for teaching problem solving, the organization of material in a presentation, and science facts devoid of deep understanding. Students could be successful with the computer activities and fail to learn curriculum objectives.

Model-It is the first and only piece of software I have found that supports the existing curriculum. It is a technology that can be productive throughout a curriculum. Initially I thought that the students should be experts before designing a model. But as I have indicated, the students' knowledge grows as the model is constructed. Modeling motivates the students to extend their understanding through added research, discussions, testing, and revising. An added benefit is their growth in their ability to write explanations and verbally to express their understanding of scientific concepts.

## REFERENCES

Kids as global scientists. University of Michigan, Ann Arbor.

Model-It. Cogito Learning Media, Inc., San Francisco, CA.

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

# Metacognitive Facilitation: An Approach to Making Scientific Inquiry Accessible to All

Barbara Y. White and John R. Frederiksen

---

## INTRODUCTION

Science can be viewed as a process of creating laws, models, and theories that enable people to predict, explain, and control the behavior of the world. Our objective in the ThinkerTools Inquiry Project has been to create an instructional approach that makes this view of understanding and doing science accessible to a wide range of students, including lower-achieving and younger students. Our hypothesis is that this objective can be achieved by facilitating the development of metacognitive knowledge and skills: students need to learn about the nature and utility of scientific models as well as the processes by which they are created, tested, and revised. More specifically, we want to help students acquire:

- ▶ self-knowledge, including awareness of what expertise they have, the forms that expertise can take, and when and why their expertise might be useful;
- ▶ self-regulatory skills, including skills for planning and monitoring such as determining goals and developing strategies for achieving those goals and then evaluating their progress to see whether their plan needs to be modified;

- ▶ self-improvement expertise, including expertise in reflecting on their knowledge and its use to determine how to improve both of these.

We believe that developing such metacognitive expertise is the key to acquiring inquiry skills and to “learning how to learn” in general. For further discussions regarding the nature of metacognition and the central role that it plays in learning, see Brown, 1987; Brown, Collins, and Duguid, 1989; Bruer, 1993; Collins and Ferguson, 1993; Nickerson, Perkins, and Smith, 1985; and Resnick, 1987.

## METACOGNITIVE FACILITATION FOR STUDENTS

We created the ThinkerTools Inquiry Curriculum to test our hypothesis about making scientific inquiry accessible to all students by focusing on the development of metacognitive expertise. In this curriculum, students engage in constructing and revising theories of force and motion. The curricular activities and materials are aimed at developing the knowledge and skills that students need to support this inquiry process. At the beginning of the curriculum, students are introduced to a metacognitive view of research, called “The Inquiry Cycle,” and a metacognitive process, called “Reflective Assessment,” in which they reflect on their inquiry (see Figure 1 and Table 1). The Inquiry Cycle consists of five steps: Question, Predict, Experiment, Model, and Apply. This cycle is repeated with each module of the curriculum and provides a goal structure that students use to guide their inquiry. The curricular activities focus on enabling students to develop the expertise needed to carry out and understand the purpose of the steps in the Inquiry Cycle, as well as to monitor and reflect on their progress as they conduct their research. This is achieved via an approach in which metacognitive knowledge and skills are learned through a process of scaffolded inquiry, reflection, and generalization. We call this process “Metacognitive Facilitation.”

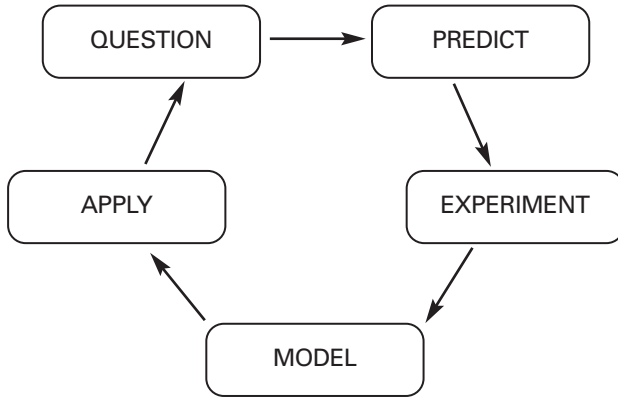
### Scaffolded Inquiry

Initially, the meaning and purpose of the steps in the Inquiry Cycle may be only partially understood by students. We therefore designed ways to provide scaffolding for their inquiry until they are able to design their own experiments and to construct their own laws to characterize their findings. We created both scaffolded activities and environments that enable students to carry out a sequence of activities corresponding to steps in the Inquiry Cycle. The scaffolded activities guide them as they do real-world experiments and help them to learn about

**FIGURE 1. A METACOGNITIVE VIEW OF THE SCIENTIFIC INQUIRY PROCESS, WHICH STUDENTS USE TO GUIDE THEIR RESEARCH**

---

THE INQUIRY CYCLE



**TABLE 1. THE CRITERIA FOR JUDGING RESEARCH THAT STUDENTS USE IN THE REFLECTIVE ASSESSMENT PROCESS**

---

CRITERIA FOR JUDGING RESEARCH

Understanding

- Understanding the Science.** Students show that they understand the science developed in the curriculum and can apply it in solving problems, in predicting and explaining real-world phenomena, and in carrying out inquiry projects.
  - Understanding the Processes of Inquiry.** Students can talk about what approach they or others have taken in exploring a research topic. For instance, they can explain what types of scientific models and inquiry processes have been used in carrying out investigations and in reaching conclusions.
  - Making Connections.** Students see the big picture and have a clear overview of their work, its purposes, and how it relates to other ideas or situations. They relate new information, ideas, and experimental results to what they already know.
-

Performance: Doing Science

- ▶ **Being Inventive.** Students are creative and examine many possibilities in their work. They show originality and inventiveness in thinking of problems to investigate, in coming up with hypotheses, in designing experiments, in creating new laws or models, and in applying their models to new situations.
  - ▶ **Being Systematic.** Students are careful, organized, and logical in planning and carrying out their work. When problems come up, they are thoughtful in examining their progress and in deciding whether to alter their approach or strategy.
  - ▶ **Using the Tools of Science.** Students use the tools and representations of science appropriately. The tools they choose to use (or create) may include such things as lab equipment, measuring instruments, diagrams, graphs, charts, calculators, and computers.
  - ▶ **Reasoning Carefully.** Students can reason appropriately and carefully using scientific concepts and models. For instance, they can argue whether or not a prediction or law that they or someone else has suggested fits with a scientific model. They can also show how experimental observations support or refute a model.
- 

Social Context of Work

- ▶ **Writing and Communicating Well.** Students clearly express their ideas to each other or to an audience through writing, diagrams, and speaking. Their communication is clear enough to allow others to understand their work and reproduce their research.
  - ▶ **Teamwork.** Students work together as a team to make progress. Students respect each others' contributions and support each others' learning. Students divide their work fairly and make sure that everyone has an important part.
-

the processes of experimental design and data analysis, the nature of scientific argument and proof, and the characteristics of scientific laws and models. The scaffolded environments include computer simulations, which allow students to create and interact with models of force and motion. They also provide analytic tools that help students analyze the results of their computer and real-world experiments. These scaffolded activities and environments make the inquiry process as easy and productive as possible at each stage in learning.

### Reflective Assessment

In conjunction with the scaffolded inquiry, students engage in a reflective process in which they evaluate their own and each other's research. This process employs a carefully chosen set of criteria, such as "Being Systematic" and "Reasoning Carefully," that characterize expert scientific inquiry (see Table 1). Students use these criteria to evaluate their work at each step in the Inquiry Cycle, which helps them to see the intellectual purpose and properties of the inquiry steps and their sequencing. Students also employ the criteria to evaluate their own and each other's work when they finish their research projects and present their work to the class. By engaging in these evaluations in which they talk about and reflect on the characteristics of expert scientific inquiry and the functions of each inquiry step, students grow to understand the nature and purpose of inquiry as well as the habits of thought that are involved.

### Generalized Inquiry and Reflection

The students employ the Inquiry Cycle and the Reflective Assessment Process repeatedly as the class addresses a series of research questions. With each repetition of the cycle, some of the scaffolding is removed so that eventually the students are conducting independent inquiry on questions of their own choosing (as in the scaffolding and fading approach of Palincsar and Brown [1984]). These repetitions of the Inquiry Cycle in conjunction with Reflective Assessment help students to refine their inquiry processes. Carrying out these activities in new research contexts also enables students to learn how to generalize the inquiry and reflection processes so that they can apply them to learning about new topics in the future.

## FACILITATING INQUIRY WITHIN THE CLASSROOM RESEARCH COMMUNITY

The project has established Classroom Research Communities in seventh, eighth, and ninth-grade science classrooms in middle schools in Berkeley and Oakland. In these classes, inquiry is the basis for developing an understanding of the physics. Physical theories are not directly taught, but are constructed by students themselves as they engage in the scaffolded inquiry and reflection. The idea is to teach students how to carry out scientific inquiry, and then have them discover the basic physical principles for themselves by doing experiments and creating theories.

The process of inquiry follows the Inquiry Cycle, shown in Figure 1, which is presented to students as a basis for organizing their explorations into the physics of force and motion. Inquiry begins with finding research questions, that is, finding situations or phenomena students do not yet understand that become new areas for investigation. Students then use their intuitions, which are often incorrect, to make conjectures about what might happen in such situations. These predictions provide them with a focus as they design experiments that allow them to observe phenomena and test their conjectures. Students then use their findings as a basis for constructing formal laws and models. By applying their models to new situations, students test the range of applicability of their models and, in so doing, identify new research questions for further inquiry.

The social organization of the research community is similar to that of an actual scientific community. Inquiry begins with a whole-class forum to develop shared research themes and areas for joint exploration. Research is then carried out in collaborative research groups. The groups thereupon reassemble to conduct a research symposium in which they present their predictions, experiments, and results, as well as the laws and causal models they propose to explain their findings. While the results and models proposed by individual groups may vary in their accuracy, in the research symposium a process of consensus building increases the reliability of the research findings. The goal is, through debate based upon evidence, to arrive at a common, agreed-upon theory of force and motion.

### Organization of the Curriculum

The curriculum is organized around a series of investigations of physical phenomena that increase in complexity. On the first day, students toss a hacky sack around the room while the teacher has them observe and list all of the factors

that may be involved in determining its motion (such as how it is thrown, gravity, air resistance, and so forth). As an inquiry strategy, the teacher suggests the need to simplify the situation, and this discussion leads to the idea of looking at simpler cases, such as that of one-dimensional motion where there is no friction or gravity (an example is a ball moving through outer space). The curriculum accordingly starts with this simple case (Module 1), and then adds successively more complicating factors such as introducing friction (Module 2), varying the mass of the ball (Module 3), exploring two-dimensional motion (Module 4), investigating the effects of gravity (Module 5), and analyzing trajectories (Module 6). At the end of the curriculum, students are presented with a variety of possible research topics to pursue (such as orbital motion or collisions), and they carry out research on topics of their own choosing (Module 7).

For each new topic in the curriculum, students follow the Inquiry Cycle:

1. **Question.** As described previously, the inquiry process begins with developing a research question such as, “What happens to the motion of an object that has been pushed or shoved when there is no friction or gravity acting on it?”
2. **Predict.** Next, to set the stage for their investigations, students try to generate alternative predictions and theories about what might happen in some specific situations related to the research question. In other words, they engage in “thought experiments.” For example, in Module 1, they are asked to predict what would happen in the following situation:

Imagine a ball that is stopped on a frictionless surface, one that is even smoother than ice. Suppose that you hit the ball with a mallet.

Then, imagine you hit the ball again in the same direction with the same size hit. Would the second hit change the velocity of the ball?

If so, describe how it would change and explain why.

In response to this question, some students might say, “the second hit does not affect the speed of the ball because it’s the same size hit as the first”; while others say, “it makes the ball go twice as fast because it gives the ball twice as much force”; and others, “it only makes the ball go a little bit faster because the ball is already moving.”

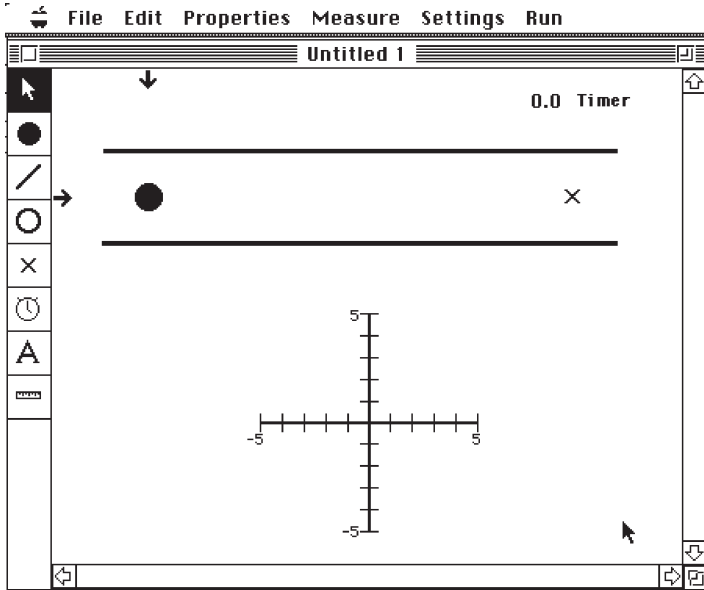
3. **Experiment.** After presenting their predictions to the class, students break into research groups to design and carry out experiments to test their alternative theories. These investigations make use of both computer simulations and real-world experimental materials.

- a. Computer activities and experiments. Computer models and experiments are made possible by use of the ThinkerTools software that we developed for the Macintosh computer. This software enables students to interact with Newtonian models of force and motion. (See Figure 2 which shows an example of a computer activity that students use in studying one-dimensional motion.) The software also lets students create their own models and experiments. Using simple drawing tools, students can construct and run computer simulations. Barriers and objects (such as the large circle shown in Figure 2) can be placed on the screen. (The objects are introduced to students as generic objects, simply called “dots,” which are the pictorial equivalent of variables that students can map onto different objects such as space ships or billiard balls.) Students can define and change the properties of any object, such as its mass, elasticity (it can be bouncy or fragile), and velocity. They can then apply impulses to the object to change its velocity using the keyboard or a joystick as in a video game. (Impulses are forces that act for a specified—usually short—amount of time like a kick or a hit.) Students can thus create and experiment with a “dot-impulse model” and can discover, for example, that applying an impulse in the same direction that the dot is moving increases the dot’s velocity by one unit of speed. In this way, they can use simulations to discover the laws of physics and their implications.

Such software enables students to create experimental situations that are difficult or impossible to create in the real world. For example, they can turn friction and gravity on or off and can select different friction laws (such as sliding friction or gas-fluid friction). They can also vary the amount of friction or gravity to see what happens. Such experimental manipulations in which students dramatically alter the parameters of the simulation make it possible for them to use inquiry strategies, such as “look at extreme cases,” which are hard to employ in real-world inquiry. This type of inquiry enables students to see more readily the behavioral implications of the laws of physics and to discover the underlying principles.

Another advantage of having students experiment with such simulations is that the software includes measurement tools that allow students to easily make accurate observations of distances, times, and velocities. These observations would often be very

**FIGURE 2. THE THINKERTOOLS SOFTWARE PROVIDES A MODELING AND INQUIRY TOOL FOR CREATING AND EXPERIMENTING WITH MODELS OF FORCE AND MOTION**



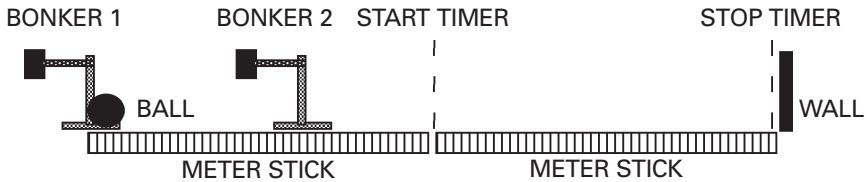
difficult to make in the corresponding real-world experiments. The software includes graphical representations of variables. As the dot moves, for example, it leaves behind “dotprints” that show how far it moved in each second and “thrustprints” that show when an impulse was applied. There is also a “datacross” that shows a dot’s  $x$  and  $y$  velocity components. And, students can have the software keep a table or graph to record, say, the velocity of the dot. In addition, there are analytic tools such as “stepping through time,” which allow students to pause the simulation and to proceed time step by time step so that they can better see and analyze what is happening to the motion of the dot. In this mode, the simulation runs for a small amount of time, leaves one dotprint on the screen, and then pauses again. The students have control over whether the simulation remains paused, proceeds to the next time step, or returns to continuous mode. These analytic tools and graphical representations help students determine the underlying laws of motion. They can also be incorporated within the students’ conceptual model to represent and reason about what might happen in successive time steps.

Ideally, the software helps students construct conceptual models that are similar to the computer's in that they both use diagrammatic representations and they both employ causal reasoning in which the computer or student steps through time to analyze events. In this way, such dynamic interactive simulations combined with these analytic tools can provide a transition from students' intuitive ways of reasoning about the world to the more abstract formal methods that scientists use for representing and reasoning about a system's behavior (White, 1993b).

- b. Real-world experiments. Students are also given a set of materials for conducting real-world experiments. This includes “bonkers” (a bonker is a rubber mallet mounted on a stand), balls of varying masses, and measurement tools such as meter sticks and stop watches (see Figure 3). These materials are coordinated with those in the ThinkerTools software. For instance, the bonker is similar to the joystick and is used to give a ball a standard-sized impulse. Employing such materials, students design and carry out real-world experiments that are related to those carried out in the simulated world. Students are also shown stop-motion videos of some of their experiments. Using frame-by-frame presentations, they can attach blank transparencies to the video screen and draw the position of a moving ball after fixed time intervals. These “dotprint analyses” allow them to measure the moment-by-moment changes in the ball's velocity.
4. **Model**. After the students have completed their experiments, they analyze their data to see whether there are any patterns. They then try to summarize and explain their findings by formulating a law and a causal model to characterize their conclusions. Students' models typically take the form: “If A then B because...”; for example, “if there are no forces like friction acting on an object, then it will go forever at the same speed, because there is nothing to slow it down.”

The computer simulations combined with real-world experiments and the process of creating a model can help students to understand the nature of scientific models. The computer is not the real world; it can only simulate real-world behavior by stepping through time and using rules to determine how forces that are acting, like friction or gravity, will change the dot's velocity on that time step. Thus, the computer is actually using a conceptual model

**FIGURE 3. AN ILLUSTRATION OF A REAL-WORLD EXPERIMENT, CALLED "THE DOUBLE-BONK EXPERIMENT," ALONG WITH THE TABLE THAT STUDENTS USE TO RECORD THEIR DATA**



**DATA TABLE**

TRIAL	DISTANCE (CM)	TIME (SEC)	VELOCITY ( $V=D \div T$ )
1	100		
2	100		
3	100		
Average	100		

to predict behavior, just as the students will use the conceptual model they construct to predict behavior. In working with the computer, the students' task is to design experiments that will help them induce the laws that are used by the simulation. This is more straightforward than the corresponding real-world inquiry task. After all, objects in the real world are not driven by laws; rather, the laws simply characterize their behavior.

One example of a modeling activity, which is carried out early in the curriculum, has students explain how their computer and real-world experiments could lead to different conclusions. They might say, for instance, that "the computer simulation does not have friction, which is affecting our real-world experiments." Alternatively, they might say that "the real world does not behave perfectly and does not follow rules." Working with a computer simulation can thus potentially help students to develop metacognitive knowledge about what scientific models are, and how laws can be used to predict and control behavior. It can also enable them to appreciate the utility of

creating computer simulations that embody scientific laws and idealized abstractions of real-world behavior, and then employing such simulations to do experiments in order to see the implications of a particular theory.

Based on the findings of their computer and real-world experiments, students prepare posters, make oral presentations to the class, and submit project reports. The Inquiry Cycle is used in organizing their reports and presentations. With writing, graphing, and drawing software (such as ClarisWorks) students analyze their data and prepare their reports. Then, in a whole-class research symposium, they evaluate together the findings from all the research groups, and choose the “best” laws and models to explain their data.

5. **Apply.** Once the class chooses the best laws and causal models, students try to apply them to different real-world situations. For instance, they might try to predict what happens when you hit a hockey puck on ice. As part of this process, they investigate the utility of their laws and models for predicting and explaining what would happen. They also investigate the limits of their models (such as, “What happens if the ice isn’t perfectly smooth?”), which inevitably raises new research questions (such as, “What are the effects of friction?”). This brings the class back to the beginning of the Inquiry Cycle and to investigating the next research question in the curriculum.

## CYCLING TOWARD INDEPENDENT INQUIRY

The Inquiry Cycle is repeated with each of the seven modules of the curriculum. As the curriculum progresses, the physics the students are dealing with increases in complexity and so does the inquiry. In the early stages of the curriculum, the inquiry process is heavily scaffolded. In Module 1, students are given experiments to do and are presented with alternative possible laws to evaluate. In this way, they see examples of experiments and laws before they have to create their own. In Module 2, students are given experiments to do but have to construct the laws for themselves. Then, in Module 3, they design their own experiments and construct their own laws to characterize their findings (see Appendix A). By the end of the curriculum, the students are carrying out independent inquiry on a topic of their own choosing.

## FACILITATING REFLECTIVE ASSESSMENT WITHIN THE CLASSROOM RESEARCH COMMUNITY

In addition to the Inquiry Cycle, which guides the students' research and helps them to understand what the research process is all about, we developed a set of criteria for characterizing good scientific research. These are presented in Table 1. They include goal-oriented criteria such as "Understanding the Science" and "Understanding the Processes of Inquiry," process-oriented criteria such as "Being Systematic" and "Reasoning Carefully," and socially-oriented criteria such as "Communicating Well" and "Teamwork." These characterizations of good work are used not only by the teachers in judging the students research projects, but also by the students themselves.

At the beginning of the curriculum, the criteria are introduced and explained to the students as the "Criteria for Judging Research" (see Table 1). Then, at the end of each step in the Inquiry Cycle, the students monitor their progress by evaluating their work on the two most relevant criteria. At the end of each module, they reflect on their work by evaluating themselves on all of the criteria. Similarly, when they present their research projects to the class, the students evaluate not only their own research projects but also each other's. They give one another feedback both verbally and in writing. These assessment criteria thus not only provide a way to introduce students to the characteristics of good research, they also help students to monitor and reflect on their inquiry processes.

In what follows, we present sample excerpts from a class's reflective assessment discussion. Students give oral presentations of their projects accompanied by a poster, and they answer questions about their research. Following each presentation, the teacher picks a few of the assessment criteria and asks students in the audience how they would rate the presentation. In these conversations, students are typically respectful of one another and generally give their peers high ratings (ratings between 3 and 5 on a five-point scale). But, within the range of high scores that they use, they do make distinctions among the criteria and offer insightful evaluations of the projects that have been presented. The following illustrates some examples of such reflective assessment conversations. (Pseudonyms are used throughout, and the transcript has been lightly edited to improve its readability.)

**Teacher:** OK, now what we are going to do is give them some feedback. What about their "understanding the process of inquiry"? In terms of their following the steps within the Inquiry Cycle, on a scale from 1 to 5, how would you score them? Vanessa?

**Vanessa:** I think I would give them a 5 because they followed everything. First they figured out what they wanted to inquire, and then they made hypotheses, and then they figured out what kind of experiment to do, and then they tried the experiment, and then they figured out what the answer really was and that Jamal’s hypothesis was correct.

**Teacher:** All right, in terms of their performance, “being inventive.” Justin?

**Justin:** Being inventive. I gave them a 5 because they had completely different experiments than almost everyone else’s I’ve seen. So, being inventive, they definitely were very inventive in their experimentation.

**Teacher:** OK, good. What about “reasoning carefully”? Jamal, how would you evaluate yourself on that?

**Jamal:** I gave myself a 5, because I had to compute the dotprints between the experiments we did on mass. So, I had to compute everything. And, I double checked all of my work.

**Teacher:** Great. OK, in terms of the social context of work, “writing and communicating well.” Carla, how did you score yourself in that area?

**Carla:** I gave myself a 4, because I always told Jamal what I thought was good or what I thought was bad, and if we should keep this part of our experiment or not. We would debate on it and finally come up with an answer.

**Teacher:** What about “teamwork”? Does anyone want to rate that? Teamwork. Nisha?

**Nisha:** I don’t know if I can say because I didn’t see them work. [laughter]

**Teacher:** That’s fine. That’s fair. You are being honest. Julia?

**Julia:** I gave them a 5 because they both talked in the presentation, and they worked together very well, and they looked out for each other.

There are various arguments for why incorporating such a Reflective Assessment Process into the curriculum should be effective. One is the “transparent assessment” argument put forward by Frederiksen and Collins (1989; Frederiksen, 1994), who argue that introducing students to the criteria by which their work will be evaluated enables students to better understand

the characteristics of good performance. In addition, there is the argument about the importance of metacognition put forward by researchers (for example, Baird et al., 1991; Brown 1987; Brown & Campione, 1996; Collins, Brown, & Newman, 1989; Miller, 1991; Reeve & Brown, 1985; Scardamalia & Bereiter, 1991; Schoenfeld, 1987; Schön, 1987; Towler & Broadfoot, 1992) who maintain that monitoring and reflecting on the process and products of one's own learning is crucial to successful learning as well as to learning how to learn. Research comparing good with poor learners shows that many students, particularly lower-achieving students, have inadequate metacognitive processes and their learning suffers accordingly (Campione, 1987; Chi et al., 1989). Thus if you introduce and support such processes in the curriculum, the students' learning and inquiry should be enhanced. Instructional trials of the ThinkerTools Inquiry Curriculum in urban classrooms (that included many lower-achieving students) provided an ideal opportunity to test these hypotheses concerning the utility of such a metacognitive Reflective Assessment Process.

#### INSTRUCTIONAL TRIALS OF THE ThinkerTools INQUIRY CURRICULUM

In 1994, we conducted instructional trials of the ThinkerTools Inquiry Curriculum. Three teachers used it in their twelve urban classes in grades seven through nine. The average amount of time they spent on the curriculum was ten and a half weeks. Two of the teachers had no prior formal physics education. They were all teaching in urban classes that averaged almost thirty students, two thirds of whom were minority students, and many were from highly disadvantaged backgrounds.

We analyzed the effects of the curriculum for students who varied in their degree of educational advantage, as measured by their standardized achievement test scores (CTBS—Comprehensive Test of Basic Skills). We compared the performance of these middle school students with that of high school physics students. We also carried out a controlled study comparing ThinkerTools classes in which students engaged in the Reflective Assessment Process with matched “control” classes in which they did not. For each of the teachers, half of the classes were reflective assessment classes and the other half were control classes. In the reflective assessment classes, the students were given the assessment framework (shown in Table 1) and they continually engaged in monitoring and evaluating their own and each other's research. In the control classes, the students were not given an explicit framework for reflecting on their research;

instead, they engaged in alternative activities in which they commented on what they did and did not like about the curriculum. In all other respects, the classes participated in the same ThinkerTools Inquiry Curriculum.

There were no significant differences in students' average CTBS scores among the classes that were randomly assigned to the different treatments (reflective assessment or control), for the classes of the three different teachers, or for the different grade levels (seventh, eighth, and ninth). (Since CTBS scores are normed for each grade level, one does not expect differences associated with grade.) Thus, the classes were all comparable with regard to achievement test scores.

## AN OVERVIEW OF THE RESULTS

Our results show that the curriculum and software modeling tools make the difficult subject of physics understandable and interesting to a wide range of students. Moreover, the emphasis on creating models enables students to learn not only about physics, but also about the properties of scientific models and the inquiry processes needed to create them. Furthermore, engaging in inquiry improves students' attitudes toward learning and doing science. Below, we provide an overview of our findings with regard to the students' development of expertise in inquiry and physics. For a more in-depth presentation of all of our results, see White & Frederiksen (1998).

### The Development of Inquiry Expertise

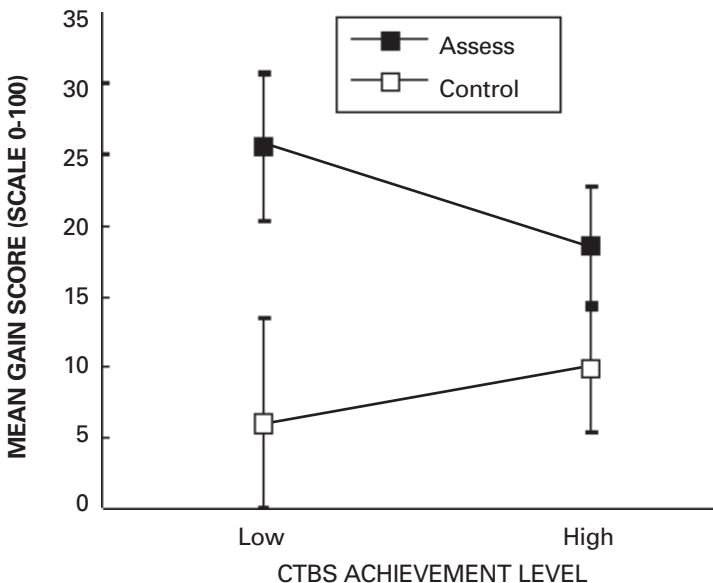
One of our assessments of students' scientific inquiry expertise was an inquiry test given both before and after the ThinkerTools Inquiry Curriculum. In this written test, the students were asked to investigate a specific research question: "What is the relationship between the weight of an object and the effect that sliding friction has on its motion?" The students were first asked to come up with alternative, competing hypotheses with regard to this question. Next, they had to design on paper an experiment that would determine what actually happens, and then they had to pretend to carry out their experiment. They had, in effect, to conduct it as a thought experiment and make up the data that they thought they would get if they actually carried out their experiment. Finally, they had to analyze their made-up data to reach a conclusion and relate this conclusion back to their original, competing hypotheses.

Scoring this test centered entirely on the students' inquiry process. Whether or not their theories embodied the correct physics was regarded as irrelevant. Figure 4 presents the gain scores on this test for both low- and high-achieving

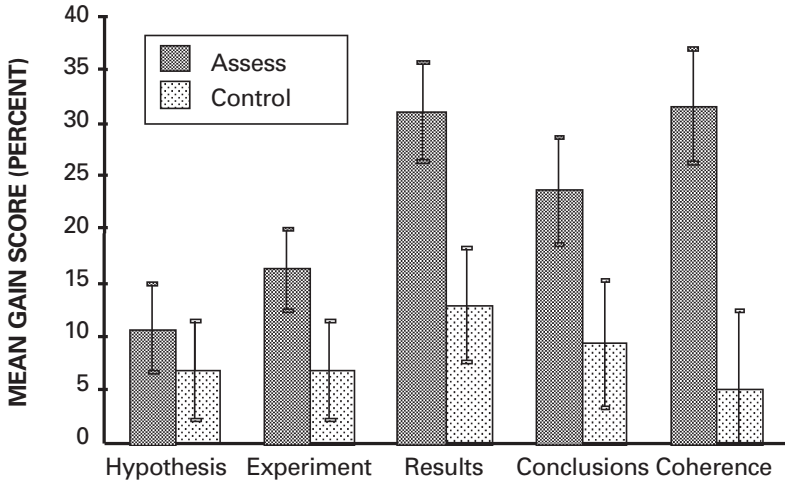
students, and for students in the reflective assessment and control classes. Notice that students in the reflective assessment classes gained more on this inquiry test, and that this was particularly true for the low-achieving students. This is the first piece of evidence that the metacognitive Reflective Assessment Process is beneficial, particularly for academically disadvantaged students.

The gain scores for each component of the inquiry test, presented in Figure 5, show that the effect of Reflective Assessment is greatest for the more difficult aspects of the test: making up results, analyzing those made-up results, and relating them back to the original hypotheses. In fact, the largest difference in the gain scores is that for a measure we call “coherence,” which assesses the extent to which the experiments that the students designed address their hypotheses, their made-up results relate to their experiments, their conclusions follow from their results, and whether they compare their conclusions with their original hypotheses. This kind of overall coherence in research is, we think, a very important indication of sophistication in inquiry. It is on this coherence measure that we see the greatest difference in favor of students who engaged in the metacognitive Reflective Assessment Process.

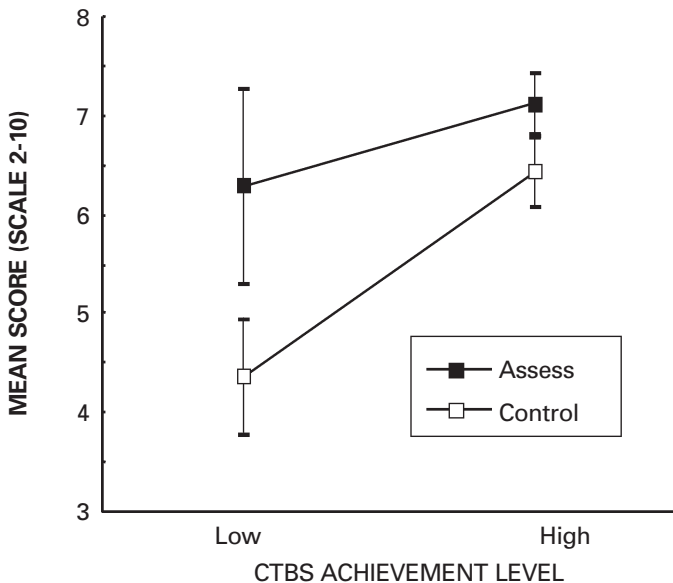
**FIGURE 4. THE MEAN GAIN SCORES ON THE INQUIRY TEST FOR STUDENTS IN THE REFLECTIVE ASSESSMENT AND CONTROL CLASSES, PLOTTED AS A FUNCTION OF THEIR ACHIEVEMENT LEVEL**



**FIGURE 5. MEAN GAINS ON THE INQUIRY TEST SUBSCORES FOR STUDENTS IN THE REFLECTIVE ASSESSMENT AND CONTROL CLASSES**



**FIGURE 6. THE MEAN SCORES ON THEIR RESEARCH PROJECTS FOR STUDENTS IN THE REFLECTIVE ASSESSMENT AND CONTROL CLASSES, PLOTTED AS A FUNCTION OF THEIR ACHIEVEMENT LEVEL**



Students carried out two research projects, one about halfway through the curriculum and one at the end. For the sake of brevity, we added the scores for these two projects together as shown in Figure 6. These results indicate that students in the reflective assessment classes do significantly better on their research projects than students in the control classes. The results also show that the Reflective Assessment Process is particularly beneficial for the low-achieving students: in the reflective assessment classes, they perform almost as well as the high-achieving students. These findings were the same across all three teachers and all three grade levels.

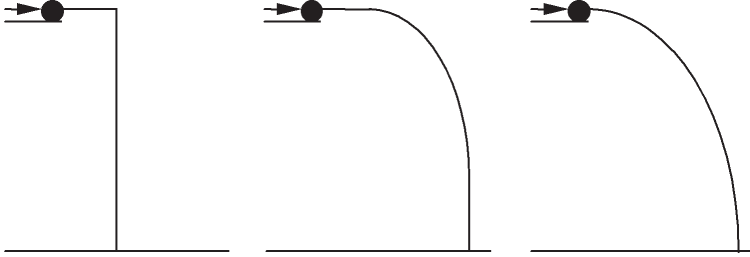
### The Development of Physics Expertise

We gave the students a General Physics Test, both before and after the ThinkerTools curriculum. This test includes items commonly used by educational researchers to assess students' understanding of Newtonian mechanics. For example, there are items such as that shown in Figure 7 in which students are asked to predict and explain how forces will affect an object's motion. On this test we found significant pretest to posttest gains. We also found that our middle school, ThinkerTools students do better on such items than do high school physics students who are taught using traditional approaches. Furthermore, on items that represent near or far transfer in relation to contexts ThinkerTools students had studied in the course, we found that there were significant learning effects on both the near and far transfer items. Together, these results show that you can teach sophisticated physics in urban, middle school classrooms when you make use of simulation tools combined with scaffolding the inquiry process. In general, this inquiry-oriented, constructivist approach appears to make physics interesting and accessible to a wider range of students than is possible with traditional approaches (White, 1993a&b; White & Frederiksen, 1998; White & Horwitz, 1988).

What is the effect of the Reflective Assessment Process on the learning of physics? The assessment criteria were chosen to address principally the process of inquiry and only indirectly the conceptual model of force and motion that students are attempting to construct in their research. Within the curriculum, moreover, students practice Reflective Assessment primarily in the context of judging their own and others' work on projects, not their progress in solving physics problems. Nonetheless, our hypothesis is that by improving the learning of inquiry skills that are instrumental in developing an understanding of physics principles, the Reflective Assessment should have an influence on students' success in developing conceptual models for the physical phenomena they have studied.

**FIGURE 7. A SAMPLE PROBLEM FROM THE PHYSICS TEST\***

Imagine that you kick a ball off a cliff.  
Circle the path the ball would take as it falls  
to the ground.



Explain the reasons for your choice:

---

---

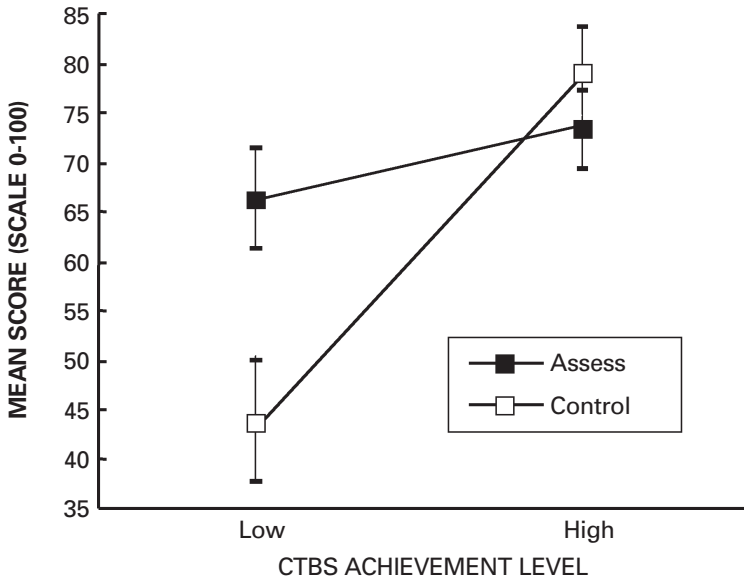
\*On a set of such items, the ThinkerTools students averaged 68% correct and significantly outperformed high school physics students who averaged 50% correct ( $t_{343} = 4.59, p = <.001$ ).

To evaluate the effects of Reflective Assessment on students' understanding of physics, we examined their performance on our Conceptual Model Test. In this test, students are asked questions about the behavior of objects in the Newtonian computer model. It assesses whether they have developed the desired Newtonian conceptual model of force and motion. Our findings, presented in Figure 8, show that for the academically disadvantaged students, the effects of Reflective Assessment extend to their learning the science content as well as the processes of scientific inquiry.

### The Impact of Understanding the Reflective Assessment Criteria

If we are to attribute these effects of introducing Reflective Assessment to students' developing metacognitive competence, we need to show that the students developed an understanding of the assessment criteria and could use them to describe multiple aspects of their work.

**FIGURE 8. THE MEAN SCORES ON THE CONCEPTUAL MODEL TEST FOR STUDENTS IN THE REFLECTIVE ASSESSMENT AND CONTROL CLASSES, PLOTTED AS A FUNCTION OF THEIR ACHIEVEMENT LEVEL**



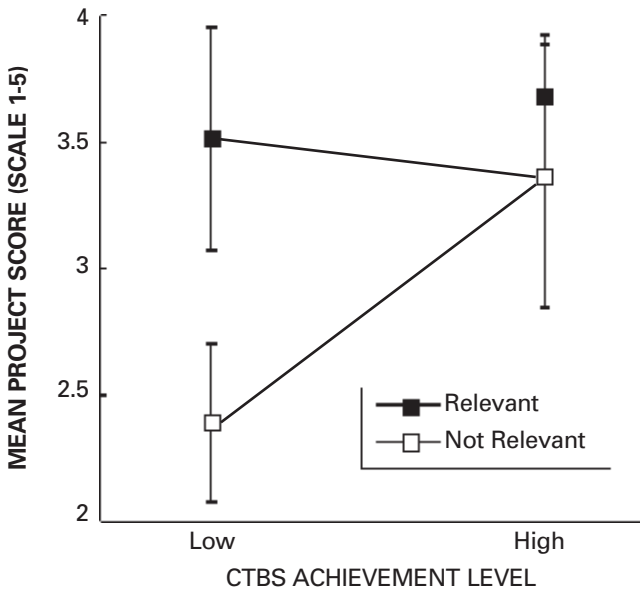
One way to evaluate their understanding of the assessment concepts is to compare their use of the criteria in rating their own work with the teachers' evaluations of their work using the same criteria. If students have learned how to employ the criteria, their self-assessment ratings should correlate with the teachers' ratings for each of the criteria. We found that students in the reflective assessment classes, who worked with the criteria throughout the curriculum, showed significant agreement with the teachers in judging their work. However, this was not the case for students in the control classes, who were given the criteria only at the end of the curriculum for use in judging their final projects. In judging Reasoning Carefully on their final projects, for instance, students in the reflective assessment classes had a correlation of .58 between their ratings and the teachers', while for students in the control classes the correlation was only .23. The average correlation for students in the reflective assessment classes over all of the criteria was .48, which is twice that for students in the control classes.

If the reflective assessment criteria are acting as metacognitive tools to help students as they ponder the functions and outcomes of their inquiry processes,

then the students' performance in developing their inquiry projects should depend upon how well they have understood the assessment concepts. To evaluate their understanding, we rated whether the evidence they cited in justifying their self assessments was relevant to the particular criterion they were considering. We then looked at the quality of the students' final projects, comparing students who had developed an understanding of the set of assessment concepts by the end of the curriculum with those who had not. Our results, shown in Figure 9, indicate that students who had learned to use the interpretive concepts appropriately in judging their work produced higher quality projects than students who had not. And again, we found that the benefit of learning to use the assessment criteria was greatest for the low-achieving students.

Taken together, these research findings clearly implicate the use of the assessment criteria as a reflective tool for learning to carry out inquiry. Students in the reflective assessment classes generated higher scoring research reports than those in the control classes. And, students who showed a clear understanding of the criteria produced higher quality investigations than those who showed

**FIGURE 9. THE MEAN SCORES ON THEIR FINAL PROJECTS FOR STUDENTS WHO DID AND DID NOT PROVIDE RELEVANT EVIDENCE WHEN JUSTIFYING THEIR SELF-ASSESSMENT SCORES, PLOTTED AS A FUNCTION OF THEIR ACHIEVEMENT LEVEL**



less understanding. Thus, there are strong beneficial effects of introducing metacognitive language to facilitate students' reflective explorations of their work in classroom conversations and in self assessment.

An important finding was that the beneficial effect of Reflective Assessment was particularly strong for the low-achieving students: The Reflective Assessment Process enabled them to gain more on the inquiry test (see Figure 4), and to perform close to the high-achieving students on their research projects (see Figure 6). The introduction of Reflective Assessment, while helpful to all, was thus closing the performance gap between the low- and high-achieving students. In fact, the Reflective Assessment Process enabled low-achieving students to perform equivalently to high-achieving students on their research projects when they did their research in collaboration with a high-achieving student. In the control classes, in contrast, the low-achieving students did not do as well as high-achieving students, regardless of whether they collaborated with a high-achieving student. Thus, there was evidence that social interactions in the reflective assessment classes—particularly between low- and high-achieving students—were important in facilitating learning (cf., Carter & Jones, 1994; Slavin, 1995; and Vygotsky, 1978).

## THE IMPLICATIONS OF OUR FINDINGS

We think that our findings have strong implications for what curricula that emphasize inquiry and metacognition can accomplish, particularly in urban school settings in which there are many academically disadvantaged students. More specifically, we argue that three important conclusions follow from our work:

- ▶ To be equitable, science curricula should incorporate reflective inquiry, and assessments of students' learning should include measures of inquiry expertise.
- ▶ Students should learn how to transfer the inquiry and reflective assessment processes to other domains so that they learn how to learn and can utilize these valuable metacognitive skills in their learning of other school subjects.
- ▶ Such an inquiry-oriented approach to education, in which the development of metacognitive knowledge and skills plays a central role, should be introduced early in the school curriculum, preferably at the elementary school level.

- 1. Science curricula should incorporate reflective inquiry and include assessments of students' inquiry expertise.** Our results suggest that, from an equity standpoint, curricular approaches can be created that are not merely equal in their value for, but actually enhance the learning of less-advantaged students. Moreover, adequately and fairly assessing the effectiveness of such curricula requires utilizing measures of inquiry expertise, such as our inquiry tests and research projects. If only subject-matter tests are used, the results can be biased against both low-achieving and female students. For instance, on the research projects, we found that low-achieving students who had the benefit of the Reflective Assessment Process did almost as well as the high-achieving students. And these results could not be attributed simply to ceiling effects. We also found that the male and female students did equally well on the inquiry tests and research projects. On the physics tests, however, the pattern of results was not comparable: males outperformed females (on both pretests and posttests) and the high-achieving students outperformed the low-achieving students (White & Frederiksen, 1998). Thus utilizing inquiry tests and research projects in addition to subject-matter tests not only plays a valuable role in facilitating the development of inquiry skills, it also produces a more comprehensive and equitable assessment of students' accomplishments in learning science.
- 2. Students should learn to transfer what they learn about inquiry and reflection to the rest of their school curriculum.** Students' work in the ThinkerTools Inquiry Curriculum and their performance on the various inquiry assessments indicate that they acquired an understanding of the Inquiry Cycle as well as the knowledge needed to carry out each of the steps in this cycle. They also learned about the forms that scientific laws, models, and theories can take and about how the development of scientific theories is related to empirical evidence. In addition, they acquired the metacognitive skills of monitoring and reflecting on their inquiry processes. Since all of science can be viewed as a process of constructing models and theories, both the Inquiry Cycle and the Reflective Assessment Process can be applied to learning and doing all areas of science, not just physics. Understanding and engaging in the Inquiry Cycle and Reflective Assessment Process should therefore benefit students in their future science courses.

In the subsequent work of ThinkerTools students, we see evidence of these benefits and transfer to new contexts. For example, eighth-grade students who did ThinkerTools in the seventh grade were asked to do research projects that used the Inquiry Cycle. They were free to choose topics other than physics. For instance, one group of students wanted to understand how listening to music affects performance on schoolwork. They did an experiment in which their classmates listened to different kinds of music while taking an arithmetic test. They wrote research reports that described how they followed the Inquiry Cycle in planning and carrying out their research, and they evaluated their own and one another's research using scoring criteria shown in Table 1. Their teacher reports that their performance on these projects was equal to or better than the performance on their ThinkerTools physics projects. Moreover, at the end of the curriculum, some students were asked if the Inquiry Cycle and Reflective Assessment Process could be used to help them learn other subjects. Many of their answers involved highly creative explanations of how it could be applied to domains such as social studies, mathematics, and English, as well as to other areas of science. An example is the following observation from a student who was discussing how the Inquiry Cycle could be useful (her statement has been edited to improve its readability):

I'm sure that a lot of things will need the Inquiry Cycle, like even things like a law court. See, they have to go through a cycle. Maybe not quite the same thing, but they have a question, like why did he do it. The predictions are like possible motives. There is no real experiment, but the equipment is like the murder weapon. The analysis and models are like what did they find out from the trial.... And so almost everything has to go through sort of a cycle like this.

Furthermore, all of the teachers attest to the benefits of both the Inquiry Cycle and the Reflective Assessment Process and have chosen to incorporate nonscience subjects. In order to make the valuable skills of inquiry, modeling, and reflection apply to other experimental sciences, such as biology, as well as to the learning of nonscience subjects, various approaches could be pursued. For instance, students could be introduced to a generalized version of the Inquiry Cycle (such as: Question, Hypothesize, Investigate, Analyze, Model, and Evaluate, which represents a minor transformation of the more experimentally oriented

Inquiry Cycle that students internalize during the ThinkerTools Inquiry Curriculum). This generalization could give students a metacognitive view of learning and inquiry that can be applied to any topic in which building predictive and explanatory models is important. In addition, the students could discuss how readily the Reflective Assessment Process, which uses the criteria shown in Table 1 (such as Making Connections, Reasoning Carefully, and Communicating Well), can be generalized to learning other topics within and beyond science. Having such explicit discussions of transfer in conjunction with explicitly using versions of the Inquiry Cycle and Reflective Assessment Process in their science and other curricula should enable students and teachers to appreciate and benefit from the power of metacognition. Investigating how such generalization and transfer can be achieved is a major concern of our current research (see White, Shimoda, & Frederiksen, 1999).

- 3. It is important to introduce inquiry-based learning and reflective assessment early in the school curriculum.** Another conclusion from our research is that the processes of reflective inquiry should be taught early. This would enable young students to develop metacognitive skills that are important components of expertise in learning. These skills should help the low-achieving students to overcome a major source of educational disadvantage. The findings from instructional trials of the ThinkerTools Inquiry Curriculum support the feasibility of achieving this goal. Students over a range of grades showed equal degrees of learning: We found no age differences in students' pretest or posttest scores on the inquiry test over grades ranging from seventh to ninth, nor did we find any age differences in students' gains on the physics tests. Extrapolating from these results suggests that inquiry-based science curricula could be introduced in earlier grades. Metz (1995) presents additional arguments concerning the feasibility and importance of teaching scientific inquiry to young students.

To meet this need, we are extending our work on the ThinkerTools Project to investigate how inquiry, modeling, and metacognition can be taught and assessed in earlier grades. As a first step, we are collaborating with sixth-grade teachers to develop an inquiry-oriented curriculum that utilizes the Inquiry Cycle, the Reflective Assessment Process, as well as other techniques for Metacognitive Facilitation. In this year-long curriculum, low-achieving students work in collaboration with high-achieving students to plan, carry out, and critically

evaluate research on a wide variety of topics across a number of disciplines. Our hope is that such an inquiry-oriented curriculum, which focuses on the development of metacognitive skills, will enable all students to learn how to learn at a young age, regardless of their prior educational advantages and disadvantages.

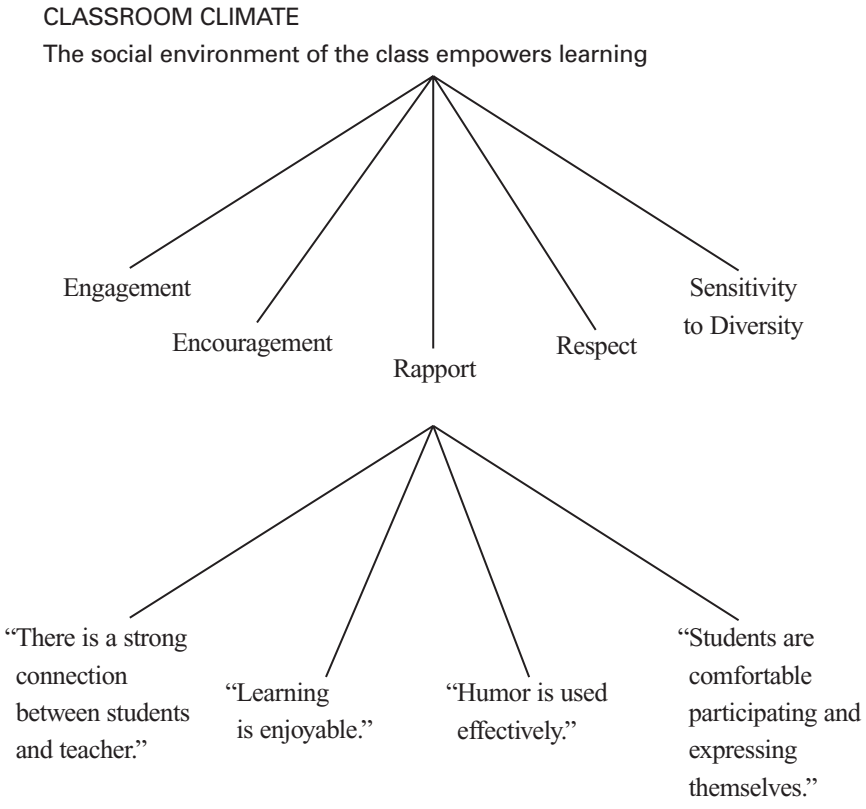
## METACOGNITIVE FACILITATION FOR TEACHERS

How can we enable teachers to implement such inquiry-oriented approaches to education? Our research in which we studied the dissemination of the ThinkerTools Inquiry Curriculum indicates that it is not sufficient to simply provide teachers with teacher's guides that attempt to outline goals, describe activities, and suggest, in a semi-procedural fashion, how the lessons might proceed (White & Frederiksen, 1998). We have found that teachers also need to develop a conceptual framework for characterizing good inquiry teaching and for reflecting on their teaching practices in the same way that students need to develop criteria for characterizing good scientific research and for reflecting on their inquiry processes.

To achieve this goal, we made use of a framework that we had developed for the National Board for Professional Teaching Standards (Frederiksen et al., 1998). This framework, which attempts to characterize expert teaching, includes five major criteria: worthwhile engagement, adept classroom management, effective pedagogy, good classroom climate, and explicit thinking about the subject matter, to which we added active inquiry. In this characterization of expert teaching, each of these criteria for good teaching is unpacked into a set of "aspects." For example, Figure 10 illustrates the criterion of "Classroom Climate," which is defined as "the social environment of the class empowers learning." Under this general criterion are five aspects: engagement, encouragement, rapport, respect, and sensitivity to diversity. Each is defined by specific characteristics of classroom practice, such as "humor is used effectively" or "there is a strong connection between students and teacher." Further, each of these is indexed to video clips, called "video snippets," which illustrate it. This framework characterizes good inquiry teaching and provides teachers with video exemplars of teaching practice.

Such materials can be used to enable teachers to learn about inquiry teaching and its value, as well as to reflect on their own and each other's teaching practices. For example, recently we tried the following approach with a group of ten student teachers. They learned to use the framework by scoring videotapes

**FIGURE 10. AN EXAMPLE OF THE HIERARCHICAL DEFINITIONS CREATED FOR EACH CRITERION, SUCH AS CLASSROOM CLIMATE, WHICH ARE USED TO CHARACTERIZE EXPERT TEACHING**



of ThinkerTools classrooms. Then, they used the framework to facilitate discussions of videotapes of their own teaching. In this way, they participated in what we call “video clubs,” which enabled them to reflect on their own teaching practices and to hopefully develop better approaches for inquiry teaching. (Video clubs incorporate social activities designed to help teachers reflectively assess and talk about their teaching practices [Frederiksen et al., 1998]). The results have been very encouraging, and our findings indicate that engaging in this reflective activity enabled the student teachers to develop a shared language for viewing and talking about teaching which, in turn, led to highly productive conversations in which they explored and reflected on their own teaching practices (Diston, 1997; Frederiksen & White, 1997; Richards & Colety, 1997).

We conclude by arguing that the same emphases on Metacognitive Facilitation that we have found is important and effective for students is beneficial for teachers as well. It can enable teachers to explore the cognitive and social goals related to inquiry teaching and to thereby improve their own teaching practices. Through this approach, both students and teachers can come to understand the goals and processes related to inquiry, and can learn how to engage in effective inquiry learning and teaching.

## ACKNOWLEDGEMENT

We gratefully acknowledge the support of our sponsors: the James S. McDonnell Foundation, the National Science Foundation, and the Educational Testing Service (ETS). The ThinkerTools Inquiry Project is a collaborative endeavor between researchers at the University of California at Berkeley and ETS and middle school teachers in the public schools of Berkeley and Oakland, California. We would like to thank all members of the team for their valuable contributions to this work.

## REFERENCES

- Baird, J., P. Fensham, R. Gunstone, and R. White. 1991. The importance of reflection in improving science teaching and learning. *Journal of Research in Science Teaching*, 28(2): 163-182.
- Brown, A. 1987. Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In *Metacognition, motivation, and understanding*, edited by F. E. Weinert and R. H. Kluwe, 60-108. Hillsdale, NJ: Erlbaum.
- Brown, A., and J. Campione. 1996. Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In *Innovations in learning: New environments for education*, edited by L. Schauble and R. Glaser, 289-325. Mahwah, NJ: Erlbaum.
- Brown, J., A. Collins, and P. Duguid. 1989. Situated cognition and the culture of learning. *Educational Researcher* 18: 32-42.

- Bruer, J. 1993. *Schools for thought: A science of learning in the classroom*. Cambridge, MA: MIT Press.
- Campione, J. 1987. Metacognitive components of instructional research with problem learners. In *Metacognition, motivation, and understanding*, edited by F. E. Weinert and R. H. Kluwe, 117-140. Hillsdale, NJ: Erlbaum.
- Carter, G., and M. Jones. 1994. Relationship between ability-paired interactions and the development of fifth graders' concepts of balance. *Journal of Research in Science Teaching* 31, 8, 847-856.
- Chi, M., M. Bassock, M. Lewis, P. Reimann, and R. Glaser. 1989. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science* 13: 145-182.
- Collins, A., J. Brown, and S. Newman. 1989. Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In *Knowing, learning, and instruction: Essays in honor of Robert Glaser*, edited by L. Resnick, 453-494. Hillsdale, NJ: Erlbaum.
- Collins, A., and W. Ferguson. 1993. Epistemic forms and epistemic games: Structures and strategies to guide inquiry. *Educational Psychologist* 28: 25-42.
- Diston, J. 1997. *Seeing teaching in video: Using an interpretative video framework to broaden pre-service teacher development*. Unpublished master's project, Graduate School of Education, University of California, Berkeley, CA.
- Frederiksen, J. (1994). Assessment as an agent of educational reform. *The Educator* 8(2): 2-7.
- Frederiksen, J., and A. Collins. 1989. A systems approach to educational testing. *Educational Researcher* 18(9): 27-32.
- Frederiksen, J., M. Sipusic, M. Sherin, and E. Wolfe. 1998. Video portfolio assessment: Creating a framework for viewing the functions of teaching. *Educational Assessment* 5(4): 225-297.

- Frederiksen, J.R., and B.Y. White. 1997. Cognitive facilitation: A method for promoting reflective collaboration. In *Proceedings of the second international conference on computer support for collaborative learning*. Toronto, Canada: University of Toronto.
- Metz, K. 1995. Reassessment of developmental constraints on children's science instruction. *Educational Researcher* 65(2): 93-127.
- Miller, M. 1991. Self-assessment as a specific strategy for teaching the gifted learning disabled. *Journal for the Education of the Gifted* 14(2):178-188.
- Nickerson, R., D. Perkins, and E. Smith. 1985. *The teaching of thinking*. Hillsdale, NJ: Erlbaum.
- Palincsar, A., and A. Brown, A. 1984. Reciprocal teaching of comprehension fostering and monitoring activities. *Cognition and Instruction* 1(2):117-175.
- Reeve, R.A., and A.L. Brown 1985. Metacognition reconsidered: implications for intervention research. *Journal of Abnormal Child Psychology* 13(3): 343-356.
- Resnick, L. 1987. *Education and learning to think*. Washington, DC: National Academy Press.
- Richards, S., and B. Colety. 1997. *Conversational analysis of the MAC-SME video analysis class: Impact on and recommendations for the MACSME program*. Unpublished master's project, Graduate School of Education, University of California, Berkeley, CA.
- Scardamalia, M., and C. Bereiter. 1991. Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences* 1(1):37-68.
- Schoenfeld, A. H. 1987. What's all the fuss about metacognition? In *Cognitive science and mathematics education*, edited by A. H. Schoenfeld, 189-215. Hillsdale, NJ: Erlbaum.

- Schön, D. 1987. *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass Publishers.
- Slavin, R. 1995. *Cooperative learning: theory, research, and practice (2nd edition)*. Needham Heights, MA: Allyn and Bacon.
- Towler, L., and P. Broadfoot. 1992. Self-assessment in primary school. *Educational Review* 44(2):137-151.
- Vygotsky, L. 1978. *Mind in society: The development of higher psychological processes*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. and Trans.). Cambridge, England: Cambridge University Press.
- White, B. 1993a. ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction* 10(1): 1-100.
- White, B. 1993b. Intermediate causal models: A missing link for successful science education? In *Advances in instructional psychology. Vol. 4.*, edited by R. Glaser, 177-252. Hillsdale, NJ: Erlbaum.
- White, B., and J. Frederiksen. 1998. Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction* 16(1): 3-118.
- White, B., and P. Horwitz. 1988. Computer microworlds and conceptual change: A new approach to science education. In *Improving learning: New perspectives*, edited by P. Ramsden, pgs. London: Kogan Page.
- White, B., T. Shimoda, and J. Frederiksen. 1999. Enabling students to construct theories of collaborative inquiry and reflective learning: Computer support for metacognitive development. *International Journal of Artificial Intelligence in Education* 10(2):121-152.

## Appendix A

This appendix contains the outline for research reports that is given to students and an example of a student's research report including data and her self-assessment.

### AN OUTLINE AND CHECKLIST FOR YOUR RESEARCH REPORTS

---

- QUESTION:**
    - ▶ Clearly state the research question.
  
  - PREDICT:**
    - ▶ What hypotheses did you have about possible answers to the question?
      - Explain the reasoning behind each of your hypotheses.
  
  - EXPERIMENT:**
    - ▶ Describe your computer experiment(s).
      - Draw a sketch of your computer model.
      - Describe how you used it to carry out your experiment(s).
    - ▶ Show your data in tables, graphs, or some other representation.
    - ▶ Describe your real-world experiment(s).
      - Draw a sketch of how you set up the lab equipment.
      - Describe how you used the equipment to carry out your experiment(s).
    - ▶ Show your data in tables, graphs, or some other representation.
  
  - MODEL:**
    - ▶ Describe how you analyzed your data and show your work.
    - ▶ Summarize your conclusions.
      - Which of your hypotheses does your data support?
      - State any laws that you discovered.
    - ▶ What is your theory about why this happens?
  
  - APPLY:**
    - ▶ Show how what you learned could be useful.
      - Give some examples.
    - ▶ What are the limitations of your investigation?
      - What remains to be learned about the relationship between the mass of an object and how forces affect its motion?
      - What further investigations would you do if you had more time?
-

**AN EXAMPLE RESEARCH REPORT ABOUT MASS AND MOTION WRITTEN BY A SEVENTH GRADE STUDENT (AGE 12)**

---

During the past few weeks, my partner and I have been creating and doing experiments and making observations about mass and motion. We had a specific question that we wanted to answer — how does the mass of a ball affect its speed?

I made some predictions about what would happen in our experiments. I thought that if we had two balls of different masses, the ball with the larger mass would travel faster, because it has more weight to roll forward with, which would help push it.

We did two types of experiments to help us answer our research question – computer and real world. For the computer experiment, we had a ball with a mass of 4 and a ball with a mass of 1. In the real world they are pretty much equal to a billiard ball and a racquetball. We gave each of the balls 5 impulses, and let them go. Each of the balls left dotprints, that showed how far they went for each time step. The ball with the mass of 4 went at a rate of 1.25 cm per time step. The ball with the mass of 1 went at a rate of 5 cm per time step, which was much faster.

For the real-world experiment, we took a billiard ball (with a mass of 166 gms) and a racquetball (with a mass of 40 gms). We bonked them once with a rubber mallet on a linoleum floor, and timed how long it took them to go 100 cm. We repeated each experiment 3 times and then averaged out the results, so our data could be more accurate. The results of the two balls were similar. The racquetball's average velocity was 200 cm per second, and the billiard ball's was 185.1 cm per second. That is not a very significant difference, because the billiard ball is about 4.25 times more massive than the racquetball.

We analyzed our data carefully. We compared the velocities, etc. of the lighter and heavier balls. For the computer experiment, we saw that the distance per time step increased by 4 (from 1.25 cm to 5 cm) when the mass of the ball decreased by 4 (from 4 to 1). This shows a direct relationship between mass and speed. It was very hard to analyze the data from our real-world experiment. One reason is that it varies a lot for each trial that we did, so it is hard to know if the conclusions we make will be accurate. We did discover that the racquetball, which was lighter, traveled faster than the billiard ball, which was heavier.

Our data doesn't support my hypothesis about mass and speed. I thought that the heavier ball would travel faster, but the lighter one always did. I did make some conclusions. From the real-world experiment I concluded that the surface of a ball plays a role in how fast it travels. This is one of the reasons that the two balls had similar velocities in our real-world experiment. (The other reason was being inaccurate.) The racquetball's surface is rubbery and made to respond to a bonk and the billiard ball's

surface is slippery and often makes it roll to one side. This made the balls travel under different circumstances, which had an effect on our results.

From the computer experiment I concluded that a ball with a smaller mass goes as many times faster than a ball with a larger mass as it is lighter than it. This happens because there is a direct relationship between mass and speed. For example, if you increase the mass of a ball then the speed it travels at will decrease.

I concluded in general, of course, that if you have two balls with different masses that the lighter one will go faster when bonked, pushed, etc. This is because the ball doesn't have as much mass holding it down.

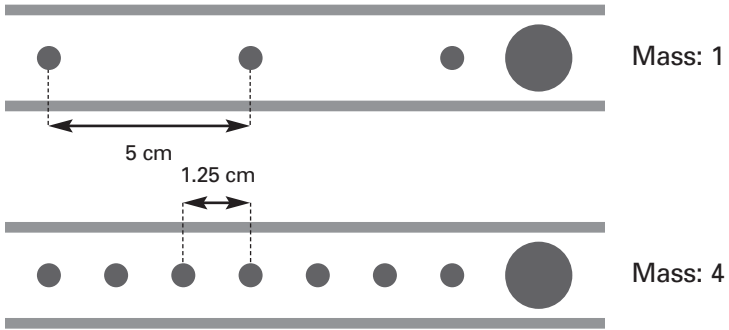
The conclusions from our experiments could be useful in real-world experiences. If you were playing baseball and you got to choose what ball to use, you would probably choose one with a rubbery surface that can be gripped, over a slippery, plastic ball. You know that the type of surface that a ball has affects how it responds to a hit. If you were trying to play catch with someone you would want to use a tennis ball rather than a billiard ball, because you know that balls with smaller masses travel faster and farther.

The investigations that we did do have limitations. In the real-world experiments the bonks that we gave the balls could have been different sizes, depending on who bonked the ball. This would affect our results and our conclusions. The experiment didn't show us how fast balls of different masses and similar surfaces travel in the real world. That is something we still can learn about. If there was more time, I would take two balls of different masses with the same kind of surface and figure out their velocities after going 100 cm.

Overall, our experiments were worthwhile. They proved an important point about how mass affects the velocity of a ball. I liked being able to come up with my own experiments and carrying them out.

---

**COMPUTER EXPERIMENTS**



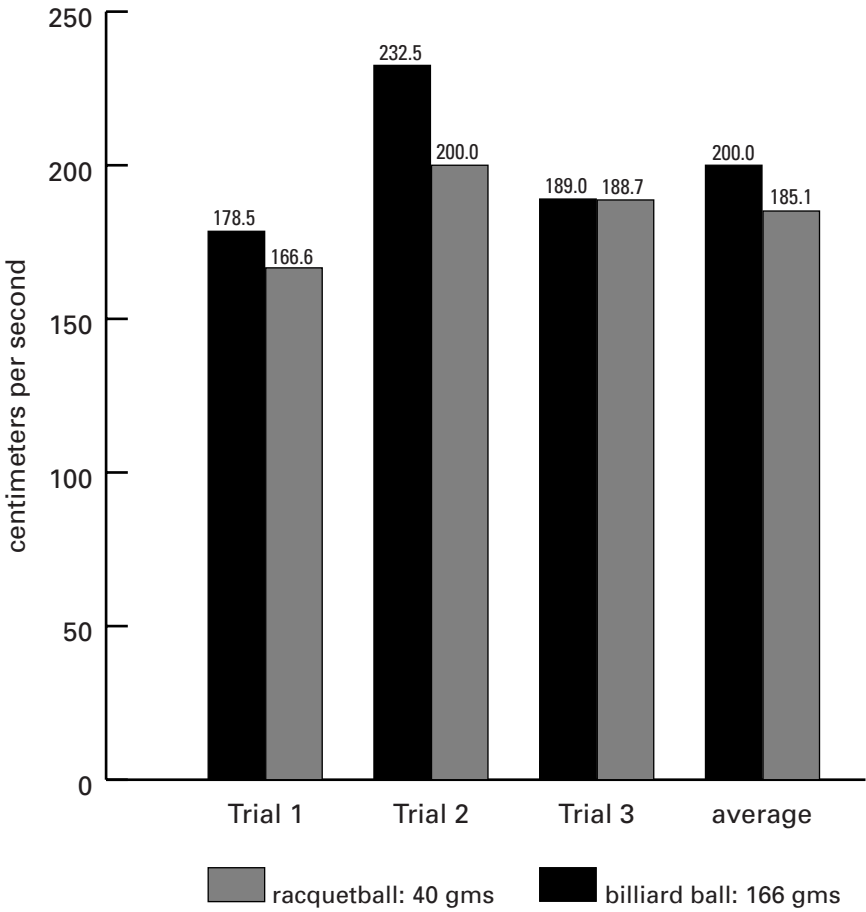
	mass	distance per time step	
	4	1.25 CM	
x 4	1	5 CM	÷ 4

<b>RACQUETBALL</b>			
mass: 40 g.	distance	time	velocity [cm per sec]
trial 1	100 cm.	.56 sec	178.5
trial 2	100 cm.	.43 sec	232.5
trial 3	100 cm.	.53 sec	189.0
average	100 cm.	.51 sec	200.0

<b>BILLIARD BALL</b>			
mass: 166 g.	distance	time	velocity [cm per sec]
trial 1	100 cm.	.60 sec	166.6
trial 2	100 cm.	.50 sec	200.0
trial 3	100 cm.	.53 sec	188.7
average	100 cm.	.54 sec	185.1

### VELOCITIES OF TWO BALLS

---



**AN EXAMPLE OF A SELF ASSESSMENT WRITTEN BY THE STUDENT WHO WROTE THE PRECEDING RESEARCH REPORT**

---

UNDERSTANDING

Understanding the Science

---

NA	1	2	3	<b>4</b>	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

---

**Justify your score based on your work.** I have a basically clear understanding of how mass affects the motion of a ball in general, but I don't have a completely clear sense of what would happen if friction, etc. was taken into account.

Understanding the Processes of Inquiry

---

NA	1	2	3	<b>4</b>	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

---

**Justify your score based on your work.** I used the inquiry cycle a lot in my write up, but not as much while I was carrying out my experiments.

Making Connections

---

NA	1	2	3	<b>4</b>	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

---

**Justify your score based on your work.** I made some references to the real world, but I haven't fully made the connection to everyday life.

## PERFORMANCE: DOING SCIENCE

## Being Inventive

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**Justify your score based on your work.** What I did was original, but many other people were original and did the same (or similar) experiment as us.

## Being Systematic

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**Justify your score based on your work.** On the whole I was organized, but if I had been more precise my results would have been a little more accurate.

## Using the Tools of Science

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**Justify your score based on your work.** I used many of the tools I had to choose from. I used them in the correct way to get results.

## Reasoning Carefully

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**Justify your score based on your work.** I took into account the surfaces of the balls in my results, but I didn't always reason carefully. I had to ask for help, but I did compute out our results mathematically.

## SOCIAL CONTEXT OF WORK

## Writing and Communicating Well

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**Justify your score based on your work.** I understand the science, but in my writing and comments I might have been unclear to others.

## Teamwork

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**Justify your score based on your work.** We got along fairly well and had a good project as a result. However, we had a few arguments.

## REFLECTION

## Self-assessment

NA	1	2	3	4	5
	NOT ADEQUATE		ADEQUATE		EXCEPTIONAL

**How well do you think you evaluated your work using this scorecard?**

I think I judged myself fairly — not too high or too low. I didn't always refer back to specific parts of my work to justify my score.

# Young Children’s Inquiry in Biology: Building the Knowledge Bases to Empower Independent Inquiry<sup>1</sup>

Kathleen E. Metz

---

## INTRODUCTION

Elementary science educators have long been concerned with achieving a “developmentally appropriate” curriculum and have for many years sought to use cognitive developmental stage theory to derive constraints on curricula at different age levels. While interpretations of Piaget’s concrete operational thought have strongly influenced science educators’ conceptualization of the inquiry processes that elementary school children *can* do, interpretations of formal operational thought have strongly influenced conceptualizations of what they can’t do.

Consider, for example, recommendations from *Science for All Children: A Guide for Improving Elementary Science Education in Your School District*, a recent publication of the National Science Resources Center:

[Piaget’s] theories still provide basic guidelines for educators about the kind of information children can understand as they move through the elementary school.... Through the primary grades, children typically group objects on the basis of one attribute, such as

color.... The significance of this information for educators is that young children are best at learning singular and linear ideas and cannot be expected to deal with more than one variable of a scientific investigation at a time.... Toward the end of elementary school, students start to make inferences.... At this stage of development, students are ready to design controlled experiments and to discover relationships among variables. (1997, pp. 28-29)

The book emphasizes the dangers of failing to teach in accordance with Piagetian stages:

If these developmental steps are not reflected in science instructional materials, there will be a mismatch between what children are capable of doing and what they are being asked to do.... When this kind of mismatch happens over and over again, children do not learn as much as they could learn about science. Equally important, they do not enjoy science. (p. 29)

The *National Science Education Standards* reflects a similar perspective:

Students should do science in ways that are within their developmental capacities.... [C]hildren in K-4 have difficulty with experimentation as a process of testing ideas and the logic of using evidence to formulate explanations.... Describing, grouping, and sorting solid objects and materials is possible early in this grade range. By grade 4, distinctions between the properties of objects and materials can be understood in specific contexts, such as a set of rocks or living materials. (National Research Council, 1996, pp. 121-123)

More generally, in line with widespread interpretations of concrete operational thought, most elementary school science curricula have emphasized the "science process skills" of observation, measurement, and organization of the concrete and, conversely, have avoided abstractions and any thinking demanding hypothetical-deductive thought—including experimental design and analysis of data sets (Metz, 1995). This approach stems from science educators' assumption that cognitive developmental stages constitute largely inflexible, hard-wired constraints on children's reasoning, within which the teacher must teach.

Nevertheless, we have negligible grounds for regarding the stages identified by cognitive developmentists as hard-wired. The reasoning that children exhibit in the developmentalist's laboratory reflects their thinking *without* the bene-

fit of instruction. Furthermore, we have robust evidence that the adequacy of individuals' reasoning is strongly impacted by the adequacy of their knowledge of the domain within which the reasoning is tested. Thus, inside the research laboratory and beyond, cognitive performance is always a complex interaction of scientific reasoning capacities and domain-specific knowledge.

Reader, think of a field in which you have negligible training—perhaps theoretical physics or evolutionary biology. If a researcher tested your reasoning capacities within this field, you could not demonstrate your best thinking and your capacities would be underestimated. In this same vein, cognitive developmental researchers focusing on preschoolers have tended to pay much more attention to using domains with which their subjects have familiarity than have the cognitive developmental researchers focusing on elementary school children. Thus, ironically, if one compares the research from these two literatures, the preschoolers frequently look *more* competent than their older peers!

As another example of the impact of knowledge on reasoning, consider the assumption—derived from interpretations of concrete operational thought—that elementary school children's thinking is bound to the concrete. Indeed, as children tend to know less than adults about most domains, their thinking is most frequently restricted to the surface, concrete features. But when by reason of particular interest or instruction children do have deeper knowledge, their thinking can be abstract. Thus, with instruction, second graders studying ecology in Brown and Campione's (in press) scientific literacy project developed a robust understanding of the abstract idea of interdependence. Conversely, adults with little knowledge of a domain can be restricted to concrete, surface features. For instance, Chi, Glaser, and Rees (1982) found that while physicists categorized physics problems on the basis of abstract physics principles, adults with little knowledge of physics categorized them on the basis of concrete features such as the kind of objects involved.

Carey has pointed out this interaction—and experimental confound—in Piaget and Inhelder's (1958) classic book, *The Development of Logical Thinking From Childhood to Adolescence*, the series of experiments upon which the elementary school science community has largely based their ideas of what elementary school children cannot do in science:

These experiments confound knowledge of particular scientific concepts with scientific reasoning more generally. It is well-documented that before the ages of 10 or 11 or so the child has not fully differentiate weight, size, and density and does not have a clear con-

ceptualization of what individuates different kinds of metals (density being an important factor). If these concepts are not completely clear in the child's mind, due to incomplete scientific knowledge, then the child will of course be unable to separate them from each other in hypothesis testing and evaluation. (Carey, 1985, p. 498)

Carey argues that in many of the domains in which Piaget and Inhelder examined the elementary school children's reasoning, the children could not make the differentiations and coordinations Piaget and Inhelder were looking for, due to their weak conceptual knowledge.

A vicious cycle has emerged here. Children's performance in the laboratory is frequently handicapped by weak knowledge of the domain within which they are tested. This weak knowledge has resulted in poor reasoning and thus an underestimation of their reasoning capacities. This underestimation of their reasoning capacities, interpreted as a ceiling on age-appropriate curricula, has resulted in unnecessarily watered-down curricula. The watered-down curriculum has led to less opportunity to learn and thus weaker domain-specific knowledge, again undermining children's scientific reasoning. In short, given the impact of adequacy of domain-specific knowledge on adequacy of scientific reasoning, this approach to defining appropriate curriculum simply maintains children's scientific reasoning more or less at the level with which they enter the classroom.

The American Association for the Advancement of Science's (AAAS) *Benchmarks for Science Literacy* (1993), while trying to delineate a frame for age-appropriate science, also acknowledges the complexity of the relation between developmentalists' descriptions of the stages of cognitive developmental and appropriate science instruction; in particular, the danger of interpreting the stages reported in the developmental literature as hard and fast limits that cannot be modified by effective instruction. *Benchmarks* states:

Research studies suggest that there are some limits on what to expect at this level of student intellectual development [grades 3rd -5th]. One limit is that the design of carefully controlled experiments is still beyond most students in the middle grades. Others are that such students confuse theory (explanation) with evidence for it and that they have difficulty making logical inferences. However, the studies say more about what students at this level do not learn in today's schools than about what they might possibly learn if instruction were more effective. (pp.10-11)

In summary, to regard the scientific reasoning that children manifest in the developmentalist's research laboratory as indicative of the level of scientific reasoning children can reflect in the classroom assumes the scientific reasoning manifested in that laboratory is largely hard-wired, an assumption that is invalidated by the profound interaction of adequacy of reasoning with adequacy of domain-specific knowledge. The interpretation of cognitive deficiencies manifested in the developmentalist's laboratory as reflections of an interaction of the adequacy of the children's knowledge with stage-based constraints demands a fundamentally different approach to curriculum design. The rest of this chapter identifies the author's conceptualization of knowledge bases needed to empower children's inquiry in biology, and then describes a curriculum aimed at scaffolding these knowledge bases and the learning the curriculum supported.

#### KNOWLEDGE BASES TO EMPOWER CHILDREN'S INQUIRY

We are working toward the elaboration of this teaching approach and examination of the scientific inquiry it supports at the elementary school level, through classroom-based teaching experiments and parallel laboratory studies. Following the lead of Ann Brown, we think of this tactic as an "educational design experiment." Brown (1992) explained, "As a design scientist in my field, I attempt to engineer innovative educational environments and simultaneously conduct experimental studies of those innovations." While the classroom context enables us to refine the learning environment and study the thinking it supports, the laboratory studies allow us to research key cognitive and instructional issues under more controlled conditions.

While there is a broad spread belief that elementary school children are incapable of independently formulating researchable questions or designing and implementing empirical studies, even at the high school level most science curriculum is not designed to foster independent inquiry (Tobin, Tippins, & Gallard, 1994). As educational psychologist Lauren Resnick (1983) noted, "we do not recognize higher-order thinking in an individual when someone else 'calls the plays' at every step" and yet this level of external control characterizes most science teaching and learning at both the elementary and high school level. In the words of *Benchmarks*:

The usual high-school science "experiment" is unlike the real thing: The question to be investigated is decided by the teacher, not the investigators; what apparatus to use, what data to collect, and how to

organize the data are also decided by the teacher (or the lab manual). (AAAS, 1993, p. 9)

Nevertheless, the gap between school science and children's science will remain immense unless we scaffold the knowledge that supports their control over the inquiry process. Narrowing the gap constitutes the top-level goal of the project.

We began with the value of maximizing the children's control over the inquiry process, stemming from our view that having significant control and responsibility over the line of inquiry constitutes a *sine qua non* of the delight and essence of doing research. At this point, given children's knowledge-base handicap, the scope of the inquiry within elementary school children's reach is unclear. We are investigating inquiry within their reach by seeking to provide instruction in the knowledge bases most important to effective data-based inquiry in biology.

The research project described in this chapter aims to develop the knowledge bases and associated metacognitive knowledge needed for the children, working in teams of two or three, to undertake largely independent scientific inquiry; conceptualizing, implementing, and revising studies in various domains of biology. We view a strong *meta* focus as crucial for this goal of educating children to take control of their own studies. Our aim is to permeate the scaffolding of the knowledge bases with a concern for metacognition (Brown, 1987), both in the sense of self-regulation and the sense of reflecting on the state of their knowledge—what they know, what they don't know.

### Scaffold Domain-Specific Knowledge

Analysis of a range of cognitive literatures leads us to conclude that if we are to support relatively powerful empirical inquiry, we need to concentrate children's science study in a relatively small number of domains. Although the idea that the depth of one's domain-specific knowledge would strongly influence the adequacy of one's reasoning seems so commonsensical that it would have failed Robert Siegler's test for a worthwhile research question (Wouldn't your grandmother have assumed that thinking is best when one has knowledge of the domain?), the idea has frequently not been reflected in children's science curricula. As Brown, Campione, Metz, and Ash (1997) stated, in the majority of children's science texts, "There is a striking lack of cumulative reference (volcanoes following magnets, following a unit on whales, etc.). This lack of coherent themes or underlying principles all but precludes systematic knowledge

building on example, analogy, principle, or theme or theory; it does not encourage sustained effort after meaning.” (pp. 20-21)

Given the deep connection between adequacy of domain-specific knowledge and adequacy of scientific reasoning, a curriculum structure which emphasizes coverage over depth handicaps inquiry. Speaking of the importance of specialization in the biologist’s thinking, evolutionary biologist, and historian of science Stephen Jay Gould writes:

No scientist can develop an adequate “feel” for nature (that undefinable prerequisite of true understanding) without probing deeply into minute empirical details of some well-chosen group of organisms. (1985, p.168).

Similarly, how can children transcend the most superficial scientific inquiry—or even be in a position to formulate questions worthy of investigation—in a domain in which they have little knowledge? If we are to support children’s thinking in the cognitively demanding task of scientific inquiry, we need to avoid the survey curriculum structure. By concentrating children’s science in a relatively small number of domains, we can support the depth of knowledge that begins to create the conditions for knowledge-building and effective scientific reasoning.

### Scaffold Knowledge of the Enterprise of Empirical Inquiry

In most of the research literature examining the impact of knowledge on scientific reasoning, “knowledge” connotes level of expertise in the field as a whole (e. g., of physics or biology) or, more specifically, one’s knowledge of that aspect of the field involved in the task at hand (e. g., Newtonian mechanics or natural selection). However, as Brewer and Samarapungavan (1991) have asserted, in their shared knowledge of the culture of science scientists draw upon knowledge that extends far beyond the bounds of their particular specialty. “Clearly,” Brewer and Samarapungavan argue, “the individual child does not yet share the enormous body of knowledge that is part of the institution of science” (p. 220). These authors identify experimental methodology and associated institutional norms as one of the fundamental knowledge bases that differentiate the scientific reasoning of adult scientists from that of children. Our project aims to scaffold the knowledge fundamental to empirical inquiry, avoiding a simplistic portrayal of some singular scientific method.

The distinction between observation and inference, for instance, undergirds the research enterprise. Thus we begin the curriculum with a focus on this distinction and at a subsequent point in the children's investigations introduce the parallel distinction between theory and evidence (and counter-evidence). Similarly, carefully designed experiments with appropriate controls are fundamental to experimental research. The curriculum scaffolds the idea of a "fair" experiment, extraneous variables, and different means of controlling them. More subtle, we are seeking to help children grasp distinctions between the norms of scientific argumentation and everyday classroom argumentation, such as the importance of recognizing and codifying uncertainty and the acceptability of a failure to reach consensus.

### Scaffold Knowledge of Domain-Specific Methodologies

A classic distinction within the cognitive science research tradition is the difference between weak and strong methods. "Weak methods" denote domain-general problem-solving strategies (such as generate and test and means/ends analysis) that one without any expertise in the domain at hand can draw upon. "Strong methods" denote the problem-solving strategies specific to the domain, and powerfully suited to its particular characteristics, purposes, and challenges. Domain experts are defined and identified in part by their knowledge of these strong methods. We conceptualize the intersection of the two knowledge bases delineated above—domain-specific knowledge and empirical inquiry—as the specialized methodological techniques of the domain under study.

To empower children's inquiry, we need to identify a repertoire of methodologies that will enable them to conduct effective investigations in the field under study. For example, while in a preliminary statistics module we focus the children's investigations on ideas of sampling, later in the study of animal behavior we teach children the domain-specific sampling techniques of time-sampling of behavior and time-sampling of location. In the study of botany, they learn other domain-specific sampling methods; including relevé sampling (continuing to double the area sampled until they study an area in which they find no additional new species) and random quadrat methods (involving random selection of which squares in the grid are to be sampled).

## Scaffold Data Representation, Data Analysis, and Fundamental Constructs of Statistics and Probability

Knowledge of how to represent and analyze data, in conjunction with fundamental constructs of statistics and probability, is fundamental for data-based scientific inquiry. We seek to develop knowledge of a repertoire of ways of representing data; including the meta-knowledge of when to use each one, what kinds of patterns each makes salient, and what each obscures. Similarly, we aim to develop elementary heuristics of data analysis, to help the children identify patterns and variability within them. Here, in particular, the heuristics vary by grade level. For example, at the second-grade level the curriculum includes analysis of the mode, median, and range, and identification of the biggest clump in the distribution of their data, whereas at the fifth-grade level children can also consider such aggregate measures as means and quartiles, and the symmetries and asymmetries in the distributions.

Even when science is taught from the perspective of active investigations, the challenges of data analysis and the uncertainty in data interpretation are rarely confronted. However, real data sets frequently are messy and doing science involves consideration of uncertainty and chance, from the point of research design to data analysis. Some fundamental knowledge of statistics and probability, at the level of ideas of chance variation, randomness, the Law of Large Numbers, and sampling, are needed for the practice of data-based research. We aim to foster these ideas through investigations focused on statistical ideas, as well as investigations focused on science domains in which children use the statistics.

### Scaffolding Knowledge of Tools

Tools such as external memory aids, calculators, data representation software, and flexible knowledge bases are considered fundamental to the scientists' work. Such tools are assumed to profoundly affect the reach of their inquiry. From the perspective of information-processing theory, these kinds of tools can fundamentally expand the processing capacities available to the human problem-solver.

Indeed, historians of science frequently identify a new tool as crucial in the development of new constructs. The computer constitutes a preeminent example of a tool fundamentally changing the bounds of human cognition along multiple parameters. In his seminal article, "The Science of Patterns,"

mathematician Lynn Steen examines the impact of the advent of computers on the field of mathematics:

[T]he computer is now the most powerful force changing the nature of mathematics. Even mathematicians who never use computers may frequently devote their entire research careers to problems generated by the presence of computers. Across all parts of mathematics, computers have posed new problems for research, provided new tools to solve old problems, and introduced new research strategies. (1988, p. 612)

In his analysis of the mathematics that made the computer possible and, in turn, the mathematics made possible by the computer, Steen documents a complex interaction of technological and conceptual advancements. Wisner and Carey's (1983) historical case study of the differentiation of heat and temperature constitutes another example of the power of tools and, as in the case of the computer, the interaction of conceptual and technological advancements. Wisner and Carey analyze changes in the understanding of heat and temperature in conjunction with an analysis of the refinement of the thermometer.

Whereas the invention of either computers or thermometers demanded significant breakthroughs, the naïve—including young children—can successfully use either tool. Thus while the thermometer supports the novice's measurement of temperature as well as the differentiation of heat and temperature (still a nontrivial distinction), the computer can support a broad range of cognitive functions. We view the scaffolding of children's knowledge of relevant tools as key to the empowering of their inquiry. For example, in our project the children learn to use visual tools such as binoculars and microscopes, in accordance with the needs of their sphere of study. They learn to use the computer for a variety of functions; including knowledge-base reference sources, simulations, vehicles to contribute to national databases, compiling their own data sets, data representation, and data analysis.

## A CURRICULUM MODULE IN ANIMAL BEHAVIOR

The author developed curriculum modules in ornithology, animal behavior, and ecology. This has been supplemented by curricula on the mathematical and statistical aspects of inquiry, including lessons drawn from the TERC curriculum *Investigations in Number, Data and Space* (e.g., Russell, Corwin, & Economopoulos, 1997; Tierney, Nemirovsky, & Weinberg; 1995) and other units developed by the author. In the first year, the fourth/fifth-grade class focused

their project-related science in ecology, while the second graders focused on animal behavior.<sup>2</sup> In the second year, the fourth/fifth-grade class studied animal behavior through a modified version of the second-grade curriculum. In the third year, second graders, third graders, and fourth/fifth graders studied ornithology in the fall and animal behavior in the spring. In the upcoming year, the foci will be botany and ecology in three classrooms, including grades one, three, and five. The three project teachers and the author are currently engaged in developing the botany curriculum.

This chapter describes how we implemented the design principles in the animal behavior module. There are a number of reasons why animal behavior constitutes a strategic domain to concentrate young children's scientific investigations. This domain of inquiry is well suited to children's analysis of patterns and variability in data, along with their reflection on ideas of uncertainty, chance, and causality. The domain is amenable to both observational and experimental research. Finally, this sphere of investigation holds high interest for children.

Children began their study of animal behavior by observing rodents in a large enclosed area; focusing on issues of behavior from a scientific perspective, scientific note taking, observation versus inference, and stimulus and response. (See Table 1.) In accordance with our *meta* focus, in those instances where there was not consensus about attribution of observation or inference, the children reported the basis of their differing attributions. This supported rich discussions of the continuum from low to high inference and the elements of interpretation that can enter in, even at this relatively basic level of the research process. Despite claims to the contrary, children even at the second-grade level, after one or two sessions, were *all* successful at the task of distinguishing inference from observation and providing appropriate justifications for their attributions. These initial observations were later elaborated by the exploration of such ideas as stimuli and response (in which the children proposed objects that might prove of interest to the rat), the challenge of trying to affect minimally what is being observed, social behavior, and the potential of variability in recorded observations between different observers.

**TABLE 1. SUMMARY OF ANIMAL BEHAVIOR CURRICULUM MODULE**

COGNITIVE OBJECTIVES	INSTRUCTIONAL ACTIVITIES	2nd	4/5th
Scientists' view of "behavior" Distinction between observation and inference (high and low).	Observe an animal, taking notes that differentiate and link observations and inferences. Collaboratively examine and critique attributions of observations and inferences.	x	x
Knowledge about the anatomy, life cycle, needs, and behavior of insects in general and crickets in particular. How is knowledge of the animal behavior used to protect and preserve species?	Directed drawings. Analysis of videos, books, knowledge-bases on computer software, and thought experiments based on scientists' accounts of resolving conceptual and/or methodological dilemmas.	x	x
Concepts of stimuli and response. Social behavior Goal of observation with minimal disturbance of the observed.	Observe and analyze behavior in response to different stimuli. Observe two animals together, focusing on social behaviors. Attention to the way in which they observe (cf. noisily) can unintentionally constitute an additional stimulus.	x	x
Conceptualization of interpretative and subjective element in doing science, even at the level of observations.	Each "research team" (mostly dyads) observes a cricket; thereafter identify- ing a behavior they both noticed but described differently; then exploring possible explanations for difference.		x
Purpose and process involved in constructing coding categories from notes regarding behavior.	Teacher helps children construct categories of behavior for time-sampling from their compiled lists of observations. Children may develop their own categories.	x	x

COGNITIVE OBJECTIVES	INSTRUCTIONAL ACTIVITIES	2nd 4/5th	
Time sampling of behavior, multiple representations of these data, with idea that different representations highlight and obscure different patterns and relationships.	Children collect time sampling of behavior data, represent in multiple ways, analyze, and formulate questions. Children enter categorical data into computer; analyze data using different representations.	x	x
Conceptualization and analysis of social behavior.	Reiterate with multiple animals to examine social behaviors.	x	x
Time sampling of location as data collection technique and analysis thereof .	Children develop maps of their crickets' terrarium. Collect and analyze time sampling of location data under solitary and social conditions; formulate questions.	x	x
Research design. Experimental controls.	Research teams select a variable for experimental manipulation and identify other variables they will need to hold constant. Collect time sampling of behavior data.	x	x
Development of experimental procedure. Representation and interpretation of numerical data.	Class collaboratively develops procedure for studying how far a cricket hops. Children enter their numerical data into computer; analyze data using different representations.	x	x
Conceptualization of a menu of research methodologies and their utility.	Class collaboratively develops a table of questions they have researched, methodologies employed, and steps involved (including methods used in prior domains).	x	x
Conceptualization of a menu of data representations and their utility.	Class collaboratively develops a table of data representations and what they are good for.		x

COGNITIVE OBJECTIVES	INSTRUCTIONAL ACTIVITIES	2nd 4/5th	
Understanding and application of the distinction between researchable and unresearchable questions.	Presented with heuristic of "Can you [begin to] answer this question through collecting data?", children brainstorm questions and discuss whether or not they are researchable.	X	X
Question formulation. Heuristics for question generation.	Throughout module children record their questions. Class explores heuristics for question generation; e.g., comparisons of same crickets under different conditions or different classes of crickets under the same conditions.	X	X
Conceptualization and implementation of a research study. Possibility of competing theories to account for same data sets.	Research teams develop question, select methodology, formulate experimental procedure, construct experimental apparatus, collect, represent and analyze data, prepare and present research poster.	X	X
Cyclical nature of science: revision of research designs and ideas. Distinction between science of discovery & science of verification.	Research teams think through additional studies needed, in order to be more confident in their findings. Research teams identify uncertainties and weaknesses in their studies and plans to address.	X	X

We made the transition from observational research to experimental research by presenting the challenge of thinking through the categories that could capture most of the behaviors on the class's compiled list of observations or, alternatively, to generate a classification schema that focuses in on the sphere(s) of behaviors on which they would personally like to concentrate their observations. Given our goal of independent inquiry, the teacher did not identify the categories, but rather presented the task to the children. As philosopher of science Karl Popper (1972) has written, the observation and classification each presume "interests, points of view, problems"; without focus on purpose, the task becomes "absurd" (p. 46). We aim to have children's categorization-construction be problem-driven. More generally, our goal here and throughout the curriculum is for children to learn how to assume responsibility and control over the decision-making process, as opposed to simply carrying out the activity as specified by the teacher.

To build toward more independence, instead of having each child in the class contribute to a compiled data set, at this point in the curriculum children were assigned to "research teams" of two children at the second-grade level and two or three at the fourth/fifth-grade level. Assignments to teams were formed largely on friendship lines, in order to facilitate smooth functioning of the collaboration. At this point where the children begin to function as research teams, we substituted crickets for the rodents, in order to provide sufficient organisms for all.

In parallel with these and subsequent empirical instructional activities, we sought to develop the children's knowledge of the anatomy, needs, life cycle, and behavior of insects in general and crickets in particular. We employed directed drawings, books the children read independently, books the teacher read to the children, videos of animal behaviorists at work, computer software, teacher presentations, and thought experiments. In our version of thought experiments, the teachers would present a question that a real animal behaviorist had faced, have the children to try to formulate strategies to address the question, and then examine the scientist's thinking. For instance, we framed one thought experiment in terms of Stephen Emlen's (1975) curiosity, theory-building, and experimentation concerning how birds found their way in the course of migration.

At first all the teams conducted the same studies, designed to promote exploration and reflection on different methodologies and different data representations. As the class moved from one research approach to the next, the teacher led the children's collaborative analysis on what the different methodological, analytical, and representational techniques elucidated and obscured.

After completion of these studies, the teacher helped the class construct two tables, one of data representations, the other of methodologies, that explicitly

defined the range of techniques they had employed (including ones from prior modules), the purpose(s) for which they had used the technique, and more generally what it was good for. We viewed these menus as external memory aids that we hoped would eventually become internalized. Through the process of constructing these “menus” of possibilities and the end-product of the two tables, we aimed to support the students’ independent decision-making in the context of their own studies.

The next phase involved considerably more independence, as the teams formulated their own research project. This phase of the design and implementation of the teams’ research studies was the most challenging and exciting to children and teacher alike, which we attribute to the fact that the question they investigate is their own and the pathway they take to investigate it of their own design. This project drew intensively on the knowledge—conceptual, procedural, epistemological, and metacognitive—that the children had developed earlier in the unit. A recurrent theme in the transfer literature is the difficulty of knowing when knowledge learned in one situation is relevant in another (Pea, 1987). Transfer becomes a particularly challenging issue in the context of experimental research, as children need to be able to decide not only what technique may be well suited to a particular question, but also how to implement or adapt the technique for the particular characteristics of the task at hand.

The first challenge involved formulation and refinement of the question for investigation. We had tried to engineer the classroom learning environments as cultures that fostered generation of questions above and beyond requests for clarification. Throughout the module the children had been recording, in their science notebooks, questions they had about cricket behavior. At this point in their study, we also explored with them different heuristics for generating researchable questions. One crucial consideration is whether a question is amenable to empirical investigation. The heuristic we used here is “Can you [begin] to answer this question through collecting data?” In this context, we also scaffolded the children’s exploration of limits in the power of a single study to resolve a question. Additional heuristics used for brainstorming questions included identifying particularly interesting behaviors or spheres thereof (for example, chirping, hopping, or aggressive behaviors); conceptualizing conditions which might affect cricket behavior; listing classes of crickets for which they had reason to anticipate variability in behavior; and, at a more aggregate level of question formation, comparing different crickets under the same condition or of the same crickets under different conditions.

Following formulation of their research question, the research teams started the process of designing their study by identifying the most appropriate methodology. While most children selected a method from the menu of possibilities, some invented approaches outside of it (for instance, planning to study the competence of crickets in working their way through a maze to get food, by building a maze and charting the crickets' progress through the maze over time). After checking back with the teacher and obtaining her agreement that their question was researchable and the methodology they had selected was well suited to the question, the children planned their study in more detail. At the second-grade level, this included a reiteration, frequently in more refined form, of the question to be investigated, a listing of materials they would need, and a step-by-step description of their experimental procedure. At the fourth/fifth-grade level, the teams' elaboration of their study included: (a) specification of the question under investigation; (b) if they conceptualized their study, not as an exploratory study, but as one testing an hypothesis, then their hypothesis and their reasoning underlying it; (c) step-by-step delineation of their experimental procedure, such that another investigator could replicate their study; (d) delineation of the extraneous variables that they thought might affect their results; (e) their plan for coping with these extraneous variables (with the caveat that it might not be feasible to control for all of them); (f) a sketch of the data sheet on which they planned to record their data; and (g) how their research plan would help them reduce the possibility that their results were simply due to chance.

The children's subsequent work—building their experimental apparatus, collecting their data, entering their data on the computer, selecting two data representations that suited their purposes—was implemented largely independently, with the teacher helping children think through occasional stumbling blocks. We structured this process through specification of the end-state, in the form of a spatial representation of a research poster and its constituent parts. Components of the poster for second graders were:

- (a) Question: Write your question here. If you had an idea about what would happen, explain what you thought would happen and why. You could say why you thought this question was interesting or how you came up with the question.
- (b) Organisms: Tell what crickets you used (what kinds, how many of each).
- (c) Materials: Make a list of materials you used.
- (d) Experimental Procedure: Describe step-by-step your procedure. Include what you kept the same and what you varied. Say

enough about how you did it, so another scientist could repeat your study.

- (e) Experimental Set-Up: Draw a picture of how you set up your materials to collect data.
- (f) Data representations: Include two representations of your data.
- (g) Data Analysis: What do your data tell you about your question? What is your evidence AND counter-evidence for your interpretation?
- (h) Conclusions: What did your study tell you about your question? What are you sure of? What are you not sure of? If you think there is more than one way of explaining your results, include both. What new questions does your study raise?"

The fourth/fifth-grade template consisted of a somewhat more elaborated version of the second-grade version. The teachers found this template very useful in scaffolding independence, as any team of children could go back to the template to see what needed to be done next.

In the last phase of the project, each team critiqued their study and designed a new one in a group interview with the author (a component we are now incorporating into the curriculum itself). The children were asked to think through how sure they were of their results and how they could be more sure. Subsequently, they are challenged to formulate changes that would improve their study and to explain why these changes would strengthen it. They were also posed the thought experiment: given the chance to do another study of cricket behavior, what question would they research and what method would they use? This activity helped the children to think further about issues connected with the science of discovery ("What would I like to find out? What am I curious about?"), as well as the science of verification ("Have I been fooled?"). These questions also encouraged the children to think more deeply about issues of research design.

## EXAMPLES OF CHILDREN'S RESEARCH PROJECTS AND THEIR REVISIONS THEREOF

Across the age span from second through fifth grade, all the teams—including those comprised of two academically weak children—formulated a researchable question and corresponding research plan, and subsequently implemented their study. Second and fifth graders alike were concerned about “fair tests”; they identified variables that could affect their results that they needed to hold constant or, where they couldn't or didn't keep them the same, needed to keep in mind as they sought to interpret their findings. All children formulated ideas of how to improve their research design that would indeed have strengthened their study. Our laboratory studies of the children's understanding of scientific inquiry indicate a much higher level of sophistication than that generally attributed to the elementary school child.

At the second-grade level, studies that pairs of children have formulated include: (a) “Do crickets act differently inside and outside?” relying first on careful observation to construct their coding categories and then time-sampling of behavior; (b) “Do crickets stay away from other bugs?” using time-sampling of location; (c) “Do crickets act differently in different colors?” using time-sampling of behavior; (d) “Do crickets do more in the night or in the day?” using time-sampling of behavior; and (e) “What size do female crickets get their ovipositor (the tube through which the female lays her eggs and deposits them in the ground)?” by measuring under a microscope (after slowing them down by putting them in the refrigerator). At the fourth/fifth-grade level, studies that pairs or triads have conducted include: (a) “What are the crickets' defenses in defending their territory and their mate?” using first careful observations to identify the defenses and subsequently time-sampling of behavior to research their relative frequency; (b) “How does a hissing roach and cricket nymph (at different times) [sic] affect a cricket's behavior?” using time-sampling of behavior; (c) “Does the gender affect how much the crickets eat?” using measurement; (d) “Where would crickets go on the playground?” using time-sampling of location; and (e) “Would two females fight over a male?” using time-sampling of behavior.<sup>3</sup> To provide a more detailed picture of the children's work, a case from the second grade and the fourth/fifth combined class is included here.

Second graders Ashley and Maria studied “What food do crickets like the most, dog food or apple?” On their poster, Ashley and Maria state that on the first day they put the same size of apple and dog food in their terrarium, and on the next two days checked to see how much food was missing from each. In their

final interview with the author, they contended that the crickets like apples more than dog food, basing their argument on their data:

**Ashley:** We found out which one they liked because they ate the apple before they ate the dog food.

**Maria:** We put the dog food and the apple in [the terrarium] and it was the same size. And then, the second day the apple was half gone. The dog food was still there.

**Ashley:** The dog food had just like two little bites.

**Maria:** Two little bites. And then the next day the apple was all gone and the dog food had three bites.

When the students are subsequently asked “Can you think of a way you could change how you did your study to make it even a better study?”, the girls have several ideas. Maria suggests manipulating the amounts of food, providing only a small amount of the seemingly less desirable food, in combination with moving to a “no choice” situation of sequential food presentation.

**Maria:** First we put the apple in, the most; and then the dog food a little bit and then we'll see if he eats it all. And if he eats it [the apple] all and the dog food only a little bit, then we'll be really really sure, because the apple of course was bigger.

**Author:** I want to understand your ideas and how your ideas right now will make it stronger than it was before.

**Maria:** If we put both [kinds of food] then maybe the cricket will choose. But if we put the dog food first, maybe it, it can't choose. So maybe it won't eat and just a few bites if it gets hungry. And then maybe when we put the apple in, it'll eat all of it. It won't get no choice. No choice.

Thus, within Maria's logic, if the cricket chooses hunger or minimal bites when presented the dog food and then consumes all the apple, they will have stronger evidence for their initial conclusions.

Ashley suggests another strategy for improving their research, based on the experimental confound she identifies between eating (and associated behavior of approaching the food) and social aversion.

**Ashley:** I've got another idea! I think we should put the dog food on one side [of the terrarium] and the apple on the other.

**Author:** How did you do it?

**Ashley:** We had them like this [indicating, with the distancing of her hands, a space of about two inches].

**Author:** Why do you think that would make it better?

**Ashley:** Because maybe they didn't want to be that close and they didn't eat that much. I think we should have put them [two kinds of food] on sides like that [far ends of the terrarium]. Because I think that the female or the male would be scared of the female to get close to them and it may scare them away from each other and they won't want to eat.

In the first implementation of the animal behavior curriculum at the second-grade level, we had not introduced the idea of experimental confound since we assumed that it would be too sophisticated an idea for these children. To our surprise, several children independently introduced it. In one context, the teacher had modeled how to think through the development of an experimental apparatus and experimental procedure, using the question of whether temperature affected the crickets' behavior. As part of this discussion, the class formulated a list of how to make the terrarium hot: heating pad, hot rock, or heat lamp. The next day when one team began to construct an experimental apparatus using a heat lamp to create hotter conditions, the author heard (and recorded) a second grader from another team advising, "Better not use the lamp. Then you'll have three things: Cold, hot, and light." Although the idea and concern of experimental confound was not shared with the class at large, it still was independently invented by two other research teams, including Ashley and Maria's. We now incorporate in the second-grade version of the curriculum the issue of extraneous variables and the importance of trying to hold constant all the variables outside of the one under investigation. The children have interpreted the idea as a way to make a "fair test" of the variable under investigation.

Kimberly and Matthew are second graders in the second year in which we implemented the animal behavior curriculum. They had the benefit of our curriculum revision, which included greater emphasis on research design, adequacy of sample size, and the use of computer tools—software encyclopedias and *Data Explorer*, a computer-based tool for recording of data and data representation. (We did not use the program's data analysis function at the second-grade level.) Kimberly and Matthew were curious about when the ovipositor first appears on the female cricket. They checked their reference sources, using the computer-based encyclopedia and classroom cricket books. When they

found no source that provided the answer, they decided on this question for their investigation.

Their research poster description of their experimental procedure includes: “4) Put nymphs under microscope. 5) Observe carefully measure size of nymph and ovipositor if it has one [sic]. 6) Keep accurate records.” To maximize the number measured, Kimberly and Matthew searched all three project classrooms for nymphs, finally measuring sixty-four individuals (ranging in length from five to eighteen millimeters). They used *Data Explorer* to enter their data. Their data representations were the data sheet itself and a scatterplot. As they collected their data, they developed an intuitive feel for the emergent relationship between nymph length and ovipositor length, as revealed in Kimberly's comment, “This one has got to be a male. If it was a female, its ovipositor would be about ten millimeters.”

In the context of the interview, Kimberly and Matthew reflect on the issue of when the ovipositor first emerges in the female, including the uncertainty of their results, the source of the uncertainty, and the strategy for addressing it:

**Author:** What did you find out?

**Kimberly:** That usually it has to be at least eight millimeters to get an ovipositor, because the littlest one that had an ovipositor was eight millimeters.

**Author:** How sure are you that a cricket has to be at least eight millimeters before it starts to grow?

**Kimberly:** Not very sure, because just on one study, that can't tell you anything about it. Because we only had one that was eight millimeters. And we measured it. And it was eight millimeters. But that was only one.

**Author:** Matthew, how sure are you that crickets have to be at least eight millimeters long before they get their ovipositor.

**Matthew:** I'm not so sure. Maybe there might be one that, probably, has seven millimeters that has one.

When the two are asked “Is a way to improve your study to be more sure that crickets have to be eight millimeters long before they get their ovipositor?” Kimberly immediately addresses the need for including more crickets in their study:

**Kimberly:** Study more nymphs. You have to look at um more different nymphs. And lots more.

**Matthew:** [Nodding agreement]

The children subsequently address the question of changes that they could make in their study to make it even better seemingly in biological engineering terms, trying to conceptualize changes in the environment which might result in earlier appearing ovipositors:

**Kimberly:** Maybe grass in the habitats. It might change it a little.

**Matthew:** Probably if we [pause]. Maybe if we add more food, they'll get theirs quicker. If you put them in hot, maybe they get an ovipositor quicker.

In the first year of the animal behavior curriculum at the fourth/ fifth-grade level, two fourth graders, Shayla and Leah, and a fifth grader, Jena, designed a study to get at the question "Do crickets react differently when they are in different colors of light?" using time-sampling of location. The team collected time-sampling data for six female crickets and six males; under conditions of red, clear, green, and black light (as produced by light bulbs of these colors in a dark closet). They found considerable inter-variability of behavior in the various light conditions (for example, "half of the crickets stayed still in the red light. Four of the six crickets stayed still in the green light").

When asked if there were ways they could be more sure of their findings, like most children at all grade levels they think extending the number of individuals tested and retesting them would help:

**Author:** Is there a way you could be more sure of your results?

**Shayla:** Yeah.

**Leah:** If we tested more crickets and tested the same crickets over again.

**Author:** And why would that help?

**Leah:** 'Cause they might do different things.

The team subsequently addresses the issue of a possible order effect; that is, that their results may have been due in part to the order in which the different colored lights were presented. To address this problem, they propose varying the order.

**Shayla:** If maybe you'd change the routine of the light colors. 'Cause in this one the whole time we did red, clear, green and black. Maybe red, green, black, clear.

**Author:** And why would you want to do that?

**Shayla:** 'Cause maybe it like maybe it would be on time. Like maybe they're like, maybe they could be hungry and you wouldn't know that and maybe they didn't feed them or something.

**Jena:** If it went like clear, red, black, green. Say if the green was last and they might be tired and they might stay still, so....

Leah is concerned that they have tested an insufficient range of colors and suggests adding more different light conditions:

**Leah:** If we use more lights.... Because it might make different if there's different colors of light. Like they might just not be different in those colors of light and the other colors of light they might react the exact same as the clear.

The girls raise the possibility that their results may be affected by where the experiment has been conducted:

**Jena:** It also might affect them, like what their surroundings are. It's like if we did it in here, they'd be surrounded by....

**Leah:** Books.

**Jena:** Books, yeah, and if we did it outside like in their normal habitat, then they'd probably act different.

**Author:** So what, so what's your suggestion there for what you might do differently to be sure, more sure?

**Jena:** Like we could do it in two different places. We could do it like in here and in a like a grass field or something.

Asked how they might improve their study, the girls have several ideas: increasing the number of organisms tested, controlling the experience of those tested, extending the range of light color conditions, constructing purer light conditions (where all stimuli in the testing area reflect the targeted light color), adding additional species of crickets, counterbalancing the order to address potential order effect, keeping constant a newly identified extraneous variable or, alternatively, redesigning the experimental procedure to study the crickets' behavior under conditions of experimental manipulation of the extraneous variable. For example:

**Author:** When scientists finish their study, they frequently think, "Hmm, how could I change my study to make it even better." It's a

good study already, but how could you make your study an even better study?

**Jena:** Test more crickets maybe.

**Author:** Why would you want to do that, Jena?

**Jena:** To be more accurate, 'cause we like tested only six and tested each cricket four times, well like five times on each light, but there's four [lights]. You might want to test, like instead of European House Crickets, you could test like, yeah, like a tree cricket or something.

**Leah:** And we had to um make sure that they weren't nymphs, 'cause we started out with grown-ups and then we get new crickets and they were all nymphs. So we had to borrow some and those ones were used to other cages. So if we got them all the same time and they were all grown-ups and we didn't have to get somebody else's. Maybe, since they have been through the same thing, . . . maybe they'd act the same.

**Jena:** 'Cause if we had just like get them out of the container thing, then maybe, and put them in the cage and tested them like a minute or a couple of minutes later, they might be exploring more. . . .

**Shayla:** They had never been in our cage before, ever. . . . Maybe, so they were kind of scared.

Conversation continues.

**Leah:** We could do an experiment on that. Are they more active in this [cage] or this [other] cage?

**Shayla:** In a different arrangement.

**Jena:** It would give us more of a variety to see the crickets. 'Cause the crickets [that the team borrowed] had never been in there and we changed the cage [introduced the novel cage] and the crickets. I think we should've maybe like tested, we should have left the cage that way and took the crickets from the first cage we had and then we should have tested them in this and then take the other four crickets [their own crickets] and test them in this cage. We should have tested all the crickets in here, that cage, and all the crickets in that cage.

**Shayla:** Because—

**Leah:** Maybe they would act different. We would have two tests, two tests on each. So if like somebody didn't know anything about our experiment, they could look at it and say, oh, they did it in two

different kinds of cages to see if they reacted differently in different colors of light and different cages and they could look at each one.

Also in the first year of implementation of the animal behavior curriculum, two fourth graders, Bryant and Douglas, researched "Will smaller or larger cages affect the crickets defensive fighting methods?" using time-sampling of behavior. The data representation on their research poster represented changes in level of intensity of aggression over time. The *y* axis indicated intensity of aggression as derived from an ordering of their coding categories; the *x* axis indicated time in one-minute increments (as coded in the time-sampling), across the four experimental conditions. The four experimental conditions consisted of: (a) large container, no female; (b) small container, no female; (c) small container with female; and (d) large container with female. A change over time graph is included for both Trial One and Trial Two. The conclusions on their research poster reflect a value on aggressive behavior and consequently disappointment with its absence:

By looking at our data, you can see that small with female got the most action. If you look at trial two, you can see we had a big disappointment. The reason we think that the fighting level went down is because they already had their territory marked. The question our experiment raises is why did the crickets fight less in trial two.

In their interview with the author, Bryant and Douglas again raise the issue of how they can account for why the intensity of the fighting decreased over time and, more generally, why crickets fight:

**Bryant:** Well, our first trial they fought a lot more. The second trial, they didn't fight as much. And we think that because they had already marked their territory.

**Douglas:** And claimed it.

**Bryant:** And claimed it so the other crickets knew to stay away. 'Cause look at our results on this other one [Trial One]. Mostly all like um walking over, no movement, climbing, like drinking....

**Douglas:** It [the graph] shows what was made up of how intense the fighting was.

**Bryant:** We put headbutting at the top because we thought that one was the most intense.... Then put kicking, wrestling.

Note that in their research poster and interview, the boys clearly differentiate between their theory and evidence.

The author then brings Bryant and Douglas back to the question they initially posed and asks them how sure they are about their results.

**Author:** So will smaller or larger cage affect the crickets' defensive fighting method?

**Bryant:** We really never answered that.

**Author:** Do you think you can answer your question?

**Bryant:** But if we— Yeah, it does affect it....

**Author:** How sure are you?

**Bryant:** We're in the middle of sure, sure, and sure.

**Douglas:** Well, I'm pretty certain, but I won't say that because we were using the same crickets on both trials.

Douglas then shifts back to the question of what the crickets were fighting over: was it simply territory or something else?

**Douglas:** So I'm thinking that if we. What I'd like to do is switch around the whole cage and make something new in it.

**Bryant:** Yeah!

**Douglas:** And see if they do the same thing. But if they don't, that um it really shows that they might um, it might not affect. Still know where it is, but it might not affect them. They might have just fought for territory.

**Author:** What do you mean, change it all around?

**Douglas:** Like move everything.

**Bryant:** Like put our big water dish in the small one and put more sticks in the little one and put different kinds of sand or something.

When probed for other ways in which they could be more sure of their conclusions, the boys get into ideas of replication, increasing the number of crickets tested, and the need to restrict their study to adults. Their rationale for using only adults reflects the key issue (and methodological correlative) of different needs at different stages of the life cycle:

**Author:** Is there another way you could be more sure of your conclusion?

**Douglas:** And conducting other experiments the same way.

**Bryant:** With different crickets. The same way but with different crickets. We couldn't find enough adult crickets. 'Cause this is mostly with adults....

**Douglas:** Nymphs won't fight because they're immature and they didn't have a need for territory. But the adults needed for their mates and stuff.

Subsequently, when asked how they might improve their study, Douglas proposes further experimental manipulations to test whether the crickets' fighting is indeed over territory.

**Author:** When scientists finish a study, they frequently think about "How could I have done the study so it would be even a better study?" How might you have done it differently to make it even better.

**Douglas:** Maybe no food, because they might be fighting over the food.

This question concerning the impetus for crickets fighting stays with the boys throughout the year. Eight months later, upon hearing that as fifth graders they will be designing new animal behavior studies, Douglas came up to the author to request that he and Bryant again work together, on the basis that they hadn't finished their research project—questions still were left unanswered. Douglas, Bryant, and Sarah (with the class now divided into triads) designed and conducted a study to research the primary stimulant of the aggression: female, food, or territory.

In parallel with our first implementation of the curriculum at the second-grade level and at the fourth/fifth-grade level, we observed child-initiated scientific reasoning above our expectations and thus in subsequent years have scaffolded, for all the fourth and fifth graders, more sophisticated research design. At the fourth/fifth grade, we now introduce more strategies for coping with extraneous variables; both holding constant variables or the experimental manipulation of variables through which one can begin to explore interactions. More generally, as an iterative process, the educational design experiment enables us to continue to refine our model of the scientific reasoning within the children's reach and, correlatively, to refine a curriculum that can more effectively empower their thinking.

## THE TEACHERS' PERSPECTIVE: VALUE AND CHALLENGES OF THIS APPROACH

Clearly, effective implementation of this way of teaching places new demands on the elementary school teacher. I interviewed the three project teachers to elicit their perspective on the value and challenges of this instructional approach. Their comments reflected the dramatic, indeed systemic, change this approach had involved for them.

One central issue they themselves had addressed in the context of project participation and also anticipated as an issue for other teachers concerned the discrepancy between the model of science reflected in this curriculum and the model of science reflected in their college science courses and prior school curricula. Conceptualizing science as a way of knowing was fundamentally new for the teachers. In the teachers' words:

We grew up thinking of science as factoids. It's a whole retraining, because people weren't taught science this way. We had to learn a new way of thinking. (Fourth/fifth-grade teacher)

There is a real simple word built into this that keeps on coming to mind: That's to do science. You're not doing much science if all you ever do is just go through the steps and procedures that someone else set up. So in the college science courses I took we were doing someone else's science. And science was things to learn about. You could learn about something. Like the wind or the weather. (Third/fourth-grade teacher; her emphases)

In the same vein, all three teachers remarked on the scaffolding of scientific inquiry, in the form of the students' design and implementation of their own studies, as a core aspect of the curriculum's power:

They're trying to get more of "discovery" science in schools. This curriculum goes beyond discovery, in that the children ask the questions to discover instead of you. It forces them to think. (Third/fourth-grade teacher)

Where they got to work on their own science projects, I was just the facilitator. It was really neat to see how they would go back to the chart and see which methodology they would use. And they knew exactly what they would use on a science question. It's a lot easier to get

involvement. They are involved! When they get to the point of doing their own research studies, when they're in charge of their own research, they can do it the whole day! (Second-grade teacher; her emphasis)

It's kid-driven rather than teacher-driven. In most other science curricula, the teacher controls everything; "Here's the graph we're going to do. Here's the materials we need." Our goal is always that they'll be able to come up with their own investigation, come up with the methodology, the materials—everything. The learning responsibility, that was a really key thing. I think the specific skills we teach them, like how to think like a scientist thinks, that's the part that's missing from canned programs. (Fourth/fifth-grade teacher; her emphasis)

The relatively large degree of the research teams' independence also raised issues of classroom organization. All of the teachers spoke of the need to foster the children's responsibility from the beginning of the year—responsibility for use of time, materials, following through on a complex task, and more generally for their own learning—to the point at which you could eventually "turn your back on groups." Related to the classroom management issue, the teachers also addressed the importance of positive social interactions. "The social interactions," the fourth/fifth-grade teacher contended, "can really destroy everything. If they don't have the background of how to work together and value each other, what do you do when three kids want to do something different in a study?"

The teachers reported continuing to struggle over how best to engineer the formation of research teams. While pairs seemed to support the most concentrated involvement and responsibility of all members, this arrangement combined with absenteeism sometimes resulted in a child working alone—which for many children proved difficult. Similarly, whereas assigning a relatively weak student to work with two stronger students tended to result in a low level of contribution on the part of the low student (despite preambles intended to encourage honoring everyone's perspective), assigning weak students to work together typically resulted in their needing much more support than other groups. Friends frequently formed productive teams, but the teachers also noted that they needed to be on the lookout for how the friends interacted over academic issues, since some friends reflected patterns of asymmetry and dominance.

Finally, the teachers all spoke of the importance of resources in making the curriculum work. For example:

It helps to have books, lots of books that give kids information.  
(Second-grade teacher; her emphasis)

A really big part for me is having the resources we need to teach. The knowledge, the stuff, you. Most teachers are alone with the textbook. For people who haven't majored in science, the support is really important. (Fourth/fifth-grade teacher)

The curriculum materials were developed for the project teachers, for the purpose of exploring the possibilities of children's relatively independent data-based inquiry in biology. We are beginning to think through how we might change the curriculum materials to make them friendly to other teachers with less direct support. These issues identified by the teachers—including shifts in vision of what science is and what it means to do science, shifts in student and teacher roles, background knowledge, the nuts and bolts of running a classroom with pairs of students assuming responsibility for their own projects—are each a critical part of the systemic change involved in teaching science this way, any of which left unattended can undermine its power.

## CONCLUSIONS

We began this project skeptical about some fairly broad spread assumptions about children's scientific reasoning abilities, together with a concern with narrowing the gap between scientists' inquiry and the inquiry of children in science lessons. We viewed maximizing the students' control over the inquiry process as crucial. Stemming from the deep connection between adequacy of knowledge and adequacy of scientific reasoning, our tactic has been to try to empower young children's relatively independent scientific inquiry by scaffolding those spheres of knowledge most fundamental to inquiry: (a) domain-specific knowledge; (b) knowledge of the enterprise of empirical inquiry; (c) domain-specific methodologies; (d) data representation, data analysis, and fundamental constructs of statistics and probability; and (e) relevant tools. The curriculum aimed to permeate the teaching of each of these knowledge components with a metacognitive perspective, involving keen attention to reflections on the adequacy of their knowledge (What do I know? What do I not know?), as well the meta knowledge needed for independent inquiry (e. g., When would I want to use this? What is this good for? How can I adapt it for different situations?).

Given the advantage of a number of relevant knowledge bases, we have found the children's inquiry to extend beyond that reflected in the developmental literature and most elementary science classrooms. In short, the project children's thinking is neither restricted to the singular, linear ideas, nor tied to the concrete. They successfully engage in the complex task of designing

controlled, albeit imperfect, experiments. Their thinking about the meaning of their data and how to improve their studies reflect inferences and hypothetical-deductive thought.

We caution science educators not to rely on the cognitive developmental literature to derive schemas of age-appropriate science curriculum. Knowledge, of various forms, can dramatically extend children's scientific inquiry. Our challenge is to further explore the science within children's reach by means of a variety of teaching experiments emphasizing different aspects at the core of thinking scientifically.

## ENDNOTES

1. The preparation of this chapter was supported by the National Science Foundation (NSF) under Grant No. REC-9618871. The ideas expressed herein are the author's and do not necessarily reflect those of the NSF.
2. Each year teachers use science kits mandated by the district for subjects outside of the life sciences.
3. We were intrigued that while most teams at the upper elementary level choose to research some aspect of social behavior, we have had only one second-grade team formulate such a project; a tendency we tentatively attribute to the increasing concerns with issues of social interactions across this age span. Curriculum elaboration of the unit now underway includes strengthening this social behavior component to empower more adequately the children's research projects in this sphere.

## REFERENCES

- American Association for the Advancement of Science. 1993. *Benchmarks for science literacy*. New York: Oxford University Press.
- Brown, A.L. 1990. Domain-specific principles affect learning and transfer in children. *Cognitive Science* 14: 107-133.
- Brown, A.L. 1992. Design experiments: Theoretical and methodological challenges in evaluating complex interventions in classroom settings. *The Journal of the Learning Sciences* 2(2): 141-178.
- Brown, A.L. 1997. Metacognition, executive control, self-regulation, and other mysterious mechanisms. In *Metacognition, motivation, and*

- understanding*, edited by R.H. Kluwe, 65-116. Hillsdale, NJ: Lawrence Erlbaum.
- Brown, A., D. Ash, M. Rutherford, K. Nakagawa, A. Gordon, and J.C. Campione. 1996. Distributed expertise in the classroom. In *Distributed cognitions*, edited by G. Solomon. New York: Cambridge University Press.
- Brown, A. L., J. Campione, K.E. Metz, and D. Ash. 1997. The development of science learning abilities in children. In *Growing up with science: Developing early understanding of science*, edited by A. Burgen and K. Härnquist, 7-40. Göteborg, Sweden: Academia Europaea.
- Brewer, W., and A. Samarapungavan. 1991. Children's theories versus scientific theories: Differences in reasoning or differences in knowledge? In *Cognition and the symbolix processes: Applied and ecological perspectives*, edited by R.R. Hoffman and D.S. Palermo. Hillsdale, NJ: Lawrence Erlbaum.
- Carey, S. 1985. Are children fundamentally different kinds of thinkers than adults? In *Thinking and learning skills*, edited by S. Chipman, J. Segal, and R. Glaser, 2. Hillsdale, NJ: Lawrence Erlbaum.
- Chi, M., R. Glaser, and E. Rees. 1982. Expertise in problem solving. In *Advances in the psychology of human intelligence*, Vol. 1, edited by R. Sternberg, 7-75. Hillsdale, NJ: Lawrence Erlbaum.
- Emlen, S. T. 1975. The stellar-orientation system of a migratory bird. *Scientific American* 233(2): 102-111.
- Gould, S. J. 1985. *The flamingo's smile: Reflections on natural history*. New York: W.W. Norton & Company.
- Inhelder, B., and J. Piaget. 1958. *The growth of logical thinking from childhood to adolescence*, translated by A. Parsons and S. Milgram. New York: Basic Books. (Original work published in 1955.)
- Metz, K.E. 1995. Re-assessment of developmental assumptions in children's science instruction. *Review of Educational Research* 65(2): 93-127.

- National Research Council. 1996. *National science education standards*. Washington, DC: National Academy Press.
- National Science Resources Center. 1997. *Science for all children: A guide to improving elementary science education in your school district*. Washington, DC: National Academy Press.
- Pea, R. 1987. Putting knowledge to use. In *Education and technology in 2020*, edited by R. Nickerson and P. Zodhiates. Hillsdale, NJ: Lawrence Erlbaum.
- Popper, K. 1972. *Conjecture and refutations: The growth of scientific knowledge*. 4th ed. London: Routledge, Kegan & Paul, Ltd.
- Resnick, L. 1983. *Education and learning to think*. Washington, DC: National Academy Press.
- Russell, S. J., R. Corwin, and K. Economopoulos. 1997. *Sorting and classifying data: Does it walk, crawl or swim?* Palo Alto: Dale Seymour Publications.
- Rutherford, F.J., and A. Ahlgren. 1990. *Science for all americans*. New York: Oxford University Press.
- Steen, L.A. 1988. The science of patterns. *Science* 240: 611-614.
- Tierney, C., R. Nemirovsky, and A.S. Weinberg. 1995. *Changes over time*. Palo Alto: Dale Seymour Publications.
- Tobin, K., D.J. Tippins, and A.J. Gallard. 1994. Research on instructional strategies for teaching science, In *Handbook of research on science teaching and learning*, edited by D. L. Gabel, 43-93. New York: MacMillan Publishing Company.
- Wiser, M., and S. Carey. 1983. When heat and temperature were one. In *Mental models*, edited by D. Gentner and A. L. Stephens. Hillsdale, NJ: Lawrence Erlbaum.

# **Inquiry Learning as Higher Order Thinking: Overcoming Cognitive Obstacles**

Anat Zohar

---

Inquiry learning is a complex activity that requires much higher order thinking and a variety of cognitive performances. But school children typically have difficulties with thinking strategies that are necessary for the practice of sound inquiry (for details, see “Literature Review” below). Ignoring such difficulties may hinder successful inquiry learning.

Science teachers are trained to invest much thought in preparing detailed lessons plans when they launch on the goal of teaching a complex science topic. Teaching experience and research findings repeatedly show that after they have gone through a unit of instruction, some students still hold on to their preconceptions. Often this happens even after much time and thought have been devoted to designing a learning sequence. Teaching scientific reasoning strategies is at least as difficult as teaching scientific concepts. But few lesson plans in science are designed specifically for the purpose of inducing a change in students' scientific reasoning strategies. Most science teachers do not devote their pedagogical skills to structure learning activities that are specifically designed to foster particular thinking skills. In fact, science teachers often do not consciously think of thinking skills as explicit educational goals (Zohar, 1999). Accordingly, in order to foster thinking we first need to consider it a distinct, explicit educational goal.

The aim of this chapter is to describe common cognitive difficulties that children encounter when they engage in inquiry learning and to suggest instructional means for coping with them. We shall start with several examples to illustrate the problem, proceed to a brief review of the relevant literature, and conclude by suggesting some practical instructional means.

## ILLUSTRATIONS OF THE PROBLEM

One illustration of the problem may be based on a “confession” recently made by a colleague, now a university professor. Over twenty years ago when he took accelerated biology in high school, his teacher introduced the class to hypothesis testing. The “if A then B...” algorithm was written neatly in his notebook. He remembers his efforts to study it by heart, failing to find any meaning in the neatly written words. Whenever he had to write a lab report, he leafed through his notebook, searching for that mysterious looking algorithm in order to be able to make a match between the particular experiment conducted in class and the correct way of writing a hypothesis. Although he wrote numerous lab reports in this way (and usually got good marks), it wasn’t until college that the mystery was solved, and he finally understood the logic of hypothesis testing.

Another illustration is taken from my current work with science teachers. Towards the end of a recent school year, a seventh-grade biology class conducted an experiment to investigate whether various parts of living organisms contain water. Flowers, stems, leaves, and a piece of meat were put into four glass containers sealed with glass covers. After heating the containers, little drops of water accumulated on all four covers. When asked about the conclusions from the experiment, several students responded by describing the experimental results: little drops of water accumulated on the glass covers. The teacher, Jane, then asked a number of questions.

1. **Teacher:** The little drops on the glass cover—what is it, a result or a conclusion?
2. **Student:** (Several students answer at the same time.) A result.
3. **Teacher:** What is the difference between a result and a conclusion?
4. **Student:** A result is what’s out there, what we could see. A conclusion is what we can learn from the result.
5. **Student:** Conclusions are like a summary of all the results.
6. **Student:** The conclusion is what you can conclude from the results.

7. **Teacher:** We can't explain a word by using the very same word. A conclusion is what you can conclude. What does it mean?
8. **Student:** A conclusion is just like a result. They are the same.
9. **Teacher:** If it is the same, why do we need two separate words?
10. **Student:** Results are like, facts... We saw the drops of water. Conclusions are our ideas, what we think.

None of the students drew the correct conclusions from this experiment—that all the parts of living organisms that were examined contained water. Some children, however, revealed an ability to explain the difference between results and conclusions (see lines 4 and 10), while others missed the difference between the two concepts at both the operational and the procedural level (see lines 5, 6, and 8).

After the lesson, I sat down with Jane to watch parts of the videotaped lesson. She said that at the beginning of the year she had taught about scientific inquiry, explaining the meaning of several concepts including “conclusions.” Conclusions were discussed in general terms and defined as “what one learns from an experiment.” Yet later on during the year, when students had to describe results and conclusions in the context of their experimentation, many were unable to complete the task successfully.

“We discussed it theoretically, in general terms,” Jane said, “but then in each experiment, they kept telling me, ‘But Jane, it’s the same thing’ [results and conclusions]. Each time anybody said so, we immediately discussed whether results and conclusions are indeed the same thing or not.... And it happened several times during the year. That’s why I was a little surprised [to see that they still found it so hard by the end of the year].”

These two illustrations have a common feature. In both cases, students rote-learned some definition that pertains to a thinking skill required for inquiry learning. The professor learned a definition by heart and could use it correctly under certain circumstances, without understanding either its meaning or how it relates to the experiments conducted in class. In Jane’s class, even students who could cite the definition of conclusions and experimental results were unable to distinguish between them in specific instances. This disability is especially striking because the teacher was sensitive to her students’ difficulty and had devoted repeated attention to that issue.

These two cases are drawn from inquiry learning but not from lessons that consisted of open inquiry. Because thinking is dependent on context and content, it may be supposed that students’ reasoning difficulties are an artifact generated

by inauthentic learning environments, and therefore do not represent genuine difficulties. Students engaging in open inquiry, however, encounter the same kind of reasoning problems. Table 1 describes a sample of problems detected in a ninth-grade classroom in an urban, middle-class school, where students were engaged in open inquiry:

**TABLE 1. EXAMPLES OF REASONING PROBLEMS DETECTED  
IN AN OPEN INQUIRY CLASSROOM**

---

1. Inadequate hypotheses

- ▶ A group of students that conducted a survey to investigate whether health food improves students' school achievements, formulated several hypotheses that were phrased as research questions: "Do sweet foods improve one's energy and concentration before classes?"; "Is a student who wakes up early and eats a healthy breakfast, more alert during school than a student who had no time for breakfast?"; "Does food improve students' general feeling?"
- ▶ A second group of students who investigated the difference between the level of vitamin C in fresh and in conserved orange juice, formulated their hypothesis as "Is the amount of vitamin C in fresh orange juice higher than in conserved juice?"
- ▶ Another frequent problem regarding hypotheses is the formulation of hypotheses that are irrelevant to the research question. The hypothesis formulated by a group of students who wanted to find out whether various periods of storage affect the level of vitamin C in orange juice was: "The level of vitamin C in the juice turns to a different state" (i.e., from liquid to solid).

2. No control of variables.

- ▶ A group of students formulated their research question as: "Does the level of vitamin C decrease when orange juice is heated to several temperatures, for various periods?" The students' experimental design included heating the orange juice up to several temperatures for different periods of time.
- ▶ Another group of students defined their research question as "Does adding sugar affect the level of vitamin C in lemon juice?" Their experimental design included manipulation of several variables at once: the amount of sugar added, the temperature and state of the juice, and the type of juice.

- ▶ A third group of students investigated the effect of several temperatures on the level of vitamin C. They made a sound experimental design, but when they carried out their plan they did not pay attention to the fact that the test tubes were heated for different periods.

### 3. Mismatch between research problem and experimental design

- ▶ Students defined the following research problem: “Does adding sugar to lemon juice affect the level of vitamin C?” But their experimental design was based on a comparison between home-made and frozen lemonade.

### 4. Problems in the processing of experimental results

(Students had trouble in translating their experimental results into graphs and charts.)

- ▶ One group who conducted a survey to investigate the frequency of vegetarians among different age groups had tried to draw a pie chart of the results. Participants did not understand that the pie represented one hundred percent, and did not know what corresponded to the whole population in their own survey.

### 5. Confusion between dependent and independent variables

- ▶ In the course of conducting a survey about diets, students tried to represent their results in a graph. They were observed while discussing which of their variables is a dependent and which an independent variable. Finally, they gave up and decided to guess (they had studied about variables before but apparently were unable to transfer their knowledge to their own research).

(These examples are taken from classrooms that used the Thinking in Science Classrooms project learning materials.)

Several studies that took place in classrooms (e. g., German et al., 1996) show us that the type of problems demonstrated in Table 1 are indeed prevalent in inquiry learning. A look into some of the theoretical studies that investigated this issue may teach us about the possible cognitive source of such problems.

## LITERATURE REVIEW

A vast literature describes studies about children's scientific thinking. For the purpose here, a distinction can be made between studies that describe deficiencies in science process skills and studies that describe students' thinking from the perspective of the relationships between experimental evidence and scientific theories and hypotheses.

Science process skills are derived from a list of activities that were used traditionally, according to a positivist paradigm of science, to describe the work of scientists. Numerous inventories include somewhat different lists of science process skills (e. g., Tamir & Lunetta, 1978; Tamir, Nussinovitz, & Friedler, 1982; Lawson, 1995). Typical items on such lists include: defining a research problem, formulating hypotheses, testing hypotheses, designing experiments (including the design of adequate controls), performing experiments, collecting data, analyzing experimental data, and drawing conclusions.

The classic work of Inhelder and Piaget (1958), as well as many studies that stemmed from their work, documented children's deficiencies in abstract thinking. These deficiencies were explained by arguing that before the ages of thirteen or fifteen, children's thinking is concrete and they are incapable of performing formal logical operations. Even after fifteen, many children are still not able to carry out such abstract thinking. These findings point to a likely source of the problems many children experience in inquiry learning. Apparently, they are incapable of employing the logical thinking necessary for successful scientific inquiry. After all, several components of what Piaget termed "logical operations" are the foundations of thinking patterns applied during inquiry. The logic of hypotheses testing is necessary for sound formulation and testing of hypotheses. The ability to manipulate variables is necessary for the design of sound experiments, including the identification of variables, the differentiation between independent and dependent variables, and the control of variables. Analysis of experimental data may also require complex procedures such as the understanding of probability and correlation or the ability to make multiple representations of data, as in coding data in tables or graphs.

Clearly, even one deficiency in any of these abilities may prevent children from drawing valid conclusions. For several decades, therefore, the predominant view among science educators was that elementary school science curricula must be constrained to activities that children of the age can handle. The activities include observing, classifying, comparing, categorizing, measuring, and the drawing of inferences on the basis of these limited activities. This view, howev-

er, has been seriously criticized. The research literature following Piaget fails to support the assumption that elementary school children cannot grasp abstract ideas. Although abstract ideas tend to be more elaborated and more prevalent in subsequent ages, some abstract ideas do emerge in the elementary school years and even earlier. Young children are capable of abstract thinking especially when they engage in inquiry that is conducted in authentic contexts (see Metz, 1995 and entries in this volume by Metz and Lehrer et al.). Moreover, the literature that has been used to derive constraints on instruction has typically been based on research that describes competencies based on alienating testing conditions, apart from instruction. More and more studies indicate that suitable instruction may bring children's thinking abilities to higher ceilings than have previously been assumed possible.

Among studies that describe students' thinking from the perspective of the relationships between experimental evidence and scientific theories is that of Kuhn, who views the coordination of theories and evidence as the heart of scientific thinking (Kuhn, Amsel, & O'Loughlin, 1988; Kuhn, 1989). Kuhn's findings suggest that children (and many lay adults) do not differentiate between theory and evidence. Instead, they meld the two into a single representation of the "way things are." When the two are discrepant, children exhibit strategies for maintaining their alignments—either adjusting the theory or adjusting the evidence by ignoring it altogether, or by attending to it in a selective, distorting manner.

Carey and her colleagues (1989; 1993) investigated the understanding of seventh-grade students regarding the nature of scientific knowledge and inquiry in contrast to proper scientific inquiry. Although there are multiple views among scientists about the nature of scientific work, they tend to agree on the basic distinction between hypotheses and the experimental evidence supporting or refining it. Scientific ideas are not simply copies of the facts, but rather distinct, constructed, and manipulable entities. The seventh-grade students, it was discovered, do not appreciate this, nor do they understand that scientists' ideas affect their experimental work and vice versa. Instead, ideas are confused with experiments, and there is no acknowledgment of the theoretical motivations behind scientific experiments. Students do not know what a hypothesis is, explaining it as an idea or a guess. They also do not know where scientists get their hypotheses from, what experiments are, why scientists perform them, and how they choose which experiment to perform. Students also do not understand when and why scientists change their ideas and what they do when they get unexpected results. Most of the seventh-grade students in this study simply saw the goal of science to be the gathering of specific facts about the world.

This brief review of the literature about children's scientific reasoning difficulties is much too short to be comprehensive. But it shows that many cases seen in classrooms are embedded in thinking difficulties that are widely documented and analyzed in theoretical cognitive studies. Several of the classroom examples described in the previous section may be caused by children's difficulty in logical thinking: the mystery of the algorithm for hypothesis testing, the lack of variable control, and the confusion between dependent and independent variables. Difficulties portrayed in other classroom examples may be in accord with the findings in the studies into children's limitations in understanding the nature of scientific experimentation. Children's prevalent inability to differentiate between experimental results and conclusions may be an instance of the difficulties Kuhn defined in distinguishing between evidence and theories, since experimental results correspond with evidence while conclusions relate to scientific theories. Students' difficulties to formulate adequate hypotheses and the mismatch found between their research questions and their experimental design, may also be related to the findings of Carey et al., who speak of the inability among young students to understand the nature of scientific knowledge and inquiry (see also Klahr & Dunbar, 1988; Sodian, Zaitchik, & Carey, 1991; and Klahr & Fay, 1993).

Although there is no consensus among researchers about the sources and developmental phases of children's difficulties in scientific reasoning, it is obvious that these difficulties will have a considerable effect on how children construct their knowledge while learning by inquiry. It is therefore essential to address them during instruction.

## INSTRUCTIONAL MEANS

### Several Approaches to Instruction

Can higher order thinking skills be taught? An accumulation of evidence indicates an affirmative answer. Many recent studies show that higher order thinking skills in general and scientific inquiry skills in particular can indeed be taught, leading to considerable gains in students' reasoning abilities.

How should higher order thinking be taught? The general literature about instruction of higher order thinking contains methods that may be adaptable to instruction of scientific inquiry skills. A primary distinction in that literature concerns the difference between the "general" and the "infusion" approaches to teaching thinking (Ennis, 1989). The general approach attempts to teach thinking abilities and dispositions generally, separately from any specific curricular

content. Thinking skills are typically taught through some content that is conscripted for the purpose of teaching reasoning—local or national political issues, problems in the school cafeteria, or some school subject—but instruction aims at general thinking skills and not deep conceptual understanding of contents. The infusion approach involves the integration of thinking into the regular school curriculum. According to this approach students learn subject matter in a deep and thought provoking way, from which general principles of thinking are made explicit.

Educators working toward instruction of inquiry skills have been embracing both methods. Friedler and Tamir (1986) taught basic concepts of scientific research to high school students by designing a module that included invitations to inquiry and a variety of exercises that lead the students gradually from simple to more complex and some highly sophisticated experiences in solving problems. An evaluation of this unit showed that students who have used this module demonstrated substantial gains in applications of inquiry skills that were measured in inquiry-oriented practical laboratory tests taken from the matriculation examinations.

Adey and colleagues Shayer and Yates (1989; 1993) applied a different approach in CASE (Cognitive Acceleration through Science Education). The project aims at improving the ability to use thinking skills across multiple subjects and topics. CASE addresses ten of Piaget's formal operations by designing a set of special lessons, replacing regular science lessons every two weeks for two years. Some of the operations are directly related to inquiry: for instance, the notion of variables, dependent and independent variables, control of variables, and probability. The CASE program is firmly grounded in the cognitive literature about children's learning. Each unit utilizes several means that are proposed for bringing about long-term effects in the general ability of learners:

- ▶ **concrete preparation**—concrete activities are used to introduce the terminology and the context in which a problem is presented;
- ▶ presentation of problems in ways that will induce a **cognitive conflict**;
- ▶ special activities to foster **metacognition**;
- ▶ explicit activities that induce the **transfer of thinking strategies** to novel situations.

CASE lessons are taught as special lessons, involving topics that are not part of the regular science curriculum. To this extent they exemplify the strategy of teaching skills in the general approach. But CASE teachers are encouraged to practice transfer activities, including specific ideas for applying the reasoning

strategies taught in CASE lessons to other parts of the curriculum. An extensive evaluation program of CASE indicated that the intervention led both to gains in Piagetian measures of cognitive development and to gains in subject-matter knowledge (in science, mathematics, and English).

A third approach to teaching scientific thinking skills is applied in the Thinking in Science Classrooms (TSC) project that supplements the regular science curriculum with learning activities designed to foster scientific reasoning skills, scientific argumentation, and knowledge of scientific concepts (Zohar, Weinberger, & Tamir, 1994; Zohar, 1996). Although fostering scientific thinking skills is among the project's explicit goals, skills are not taught as context-free entities. The TSC approach is based on the assumption that scientific reasoning cannot be taught by developing discrete, decontextualized skills but must always be deeply embedded in specific contents. The contents of the learning activities always match topics from the regular science syllabus. Therefore, teachers may incorporate these activities in the course of instruction whenever they get to a topic that is covered by one of the activities. The project produces a set of opportunities, calling for "thinking events" to take place in multiple scientific topics. Thus, the project is designed according to the infusion approach to teaching higher order thinking (Ennis, 1989).

Instruction always begins with concrete problems regarding specific scientific phenomena that students are asked to solve. During the learning process, students are active. Much of their work takes place in small groups with rich scientific argumentation. After students have used the same reasoning skill in various concrete contexts, they are encouraged (usually through class discussion) to engage in metacognitive activities that include generalization, identification of skills, and formulation of rules regarding those skills. Learning of reasoning skills is therefore achieved through an inductive process in which generalizations are made by the learners themselves. In order to avoid fixed patterns of learning activities which might eventually train students to deal with them in a merely algorithmic way, varied types of learning activities were designed:

- ▶ inquiry and critical thinking skills learning activities;
- ▶ investigation of microworlds;
- ▶ learning activities promoting argumentation skills about bioethical dilemmas in genetics;
- ▶ open-ended inquiry learning activities.

Evaluation studies investigating the effect of the first, second, and third types of learning activities indicate gains in both reasoning skills and scientific knowl-

edge (Zohar, Weinberger, & Tamir, 1994; Zohar, 1996; Zohar & Nemet, submitted). An evaluation study investigating the effect of the fourth type of learning activities is currently under progress.

### Teaching Inquiry Skills: Detailed Description of One Unit

To illustrate how inquiry skills are taught as a distinct educational objective, let us turn in some detail to one of the TSC learning activities, Investigation of Microworlds.

Description of Task and Typical Students' Performance. The learning activities for investigation of microworlds consist of a computerized simulation, a set of worksheets that students employ individually or in small groups, and an instructional sequence that is taught to a whole class (Zohar, 1994, 1996). Several similar microworlds were developed in various biological topics such as photosynthesis, plant germination, ecology, and nutrition, so that teachers can choose to use a microworld that matches a particular topic they teach. The idea for these learning activities originated in tasks used in a set of theoretical studies designed to investigate the development of scientific reasoning skills (Schauble, 1990; Kuhn, Schauble, & Garcia-Mila, 1992). Rapid and universal progress in thinking skills were observed in these studies, suggesting that the tasks they used might be applied to practical educational purposes (Zohar, 1994). In order to adapt tasks from research purposes to classroom use, several major changes were introduced. The topics of the tasks were changed to match topics that are part of the science curricula and means for class management—among them worksheets and computerized databases—were added.

In a learning activity related to photosynthesis, for example, students are asked to investigate five variables—light intensity, temperature, species of plant, natural growth area, and carbon dioxide—to determine which of them affect the rate of photosynthesis as measured by the amount of oxygen released in a fixed period. Students' investigation consists of defining the variables they wish to investigate, planning a combination of variables they want to examine, conducting the simulated experiment on the computer, making inferences and justifying them.

When students begin their inquiry, their investigation is often unsystematic, characterized by the use of invalid reasoning strategies such as: ignoring evidence, conducting experiments without controlling variables, or constantly changing the focus of their inquiry. An unsystematic investigation is illustrated by Alice, an eighth-grade student, in her first few experiments with the photosynthesis problem.

Attempting to investigate the effect of light intensity, Alice conducted these two experiments:

VARIABLE	EXPERIMENT #1	EXPERIMENT #2
light intensity	1 light bulb	2 light bulbs
temperature	17 C	25 C
carbon dioxide	added	added
species of plant	R	R
natural growth area	Y	Y
Results	1/4 test tube of oxygen	3/4 test tube of oxygen

After completing the second experiment, the following exchange took place between Alice and the experimenter :

**Experimenter:** So what did you find out?

**Alice:** That adding more light is good. It makes more photosynthesis.

**Experimenter:** How do you know that?

**Alice:** Because of what I did. When I added another light bulb, there was more oxygen.

**Experimenter:** Did you find out anything else?

**Alice:** Yeah. Adding more temperature is also good.

**Experimenter:** And how do you know that?

**Alice:** Because I saw what happened when the temperature was 25. There was more oxygen.

Alice's inferences are clearly invalid because she did not control variables. In the second experiment she changed both the light intensity and the temperature. Alice's confusions are by no means exceptional. An evaluation study indicated that approximately 90% of the inferences made by junior high school students when they start investigating the microworlds are invalid. A special learning sequence was designed to advance their thinking.

Description and Rationale of Learning Sequence. The sequence consists of three stages. At the first, students conduct independent investigations of the problem. The second includes a whole class discussion structured around the issue of variable control. At the third stage, students once again resume their independent investigations of the microworld.

*Stage One.* The main goal of Stage One is to expose the thinking strategies that students employ before instruction. Since the microworld represents a simplification of an actual experiment, the results of students' unsystematic experimentation often lead to a cognitive disequilibrium, undermining the confidence in their initial unsound strategies for solving the problem (Zohar, 1996). A conversation that took place in the classroom between George, a ninth-grade student and an experimenter, who acted as a passive observer, illustrates the point.

During a lesson, George approached the experimenter with a puzzled look on his face.

**George:** I must ask you a question. I have a problem. I don't know whether what I'm proving is correct.

**Experimenter:** ??????

**George:** Look at what I just did.

(George shows his work sheet, pointing to two experiments in which he failed to control variables. He changed the levels of both light and temperature: in the first experiment he used two light bulbs and a temperature of seventeen degrees Celsius while in the second experiment he used one light bulb and a temperature of twenty-five degrees. The resulting amount of oxygen was the same in both experiments—half a tube of oxygen, because the two variables were compensating for each other.)

**George:** These results show me that both the temperature and the light make no difference. But I'm not really sure.... Because it may be both. It may be both the temperature and the light. Because both of them are being changed. So I don't know what to do.

An examination of this student's work sheet showed that during his initial experiments he had concluded that increasing the level of either light or temperature, increases the rate of photosynthesis. This student's initial experiments and inferences, then, were similar to Alice's. Although these conclusions are correct, the student drew them through the use of invalid reasoning strategies—he did not control variables. But when variables are not controlled, students often draw contradictory inferences. Indeed, this is what George did. Although in his initial experiments he had concluded that both the light and the temperature affect the rate of photosynthesis, he now concluded the opposite. The conflict with his initial findings induced a state of cognitive dissonance in his mind. As the final two lines in the excerpt show, he is now unsure about his findings and puzzled about how to carry on his investigation.

When George resumed his independent investigation, he once again changed his mind. This time he concluded that the light does make a difference but the temperature does not. Nevertheless, he was still not satisfied and rightly so, because his conclusion regarding the temperature is incorrect. He therefore concluded by saying that he would have to continue his experimentation. According to conceptual change learning theory, an undermining of one's initial thinking is an important stage in learning. It creates dissatisfaction with existing reasoning strategies and induces the motivation for adopting alternative strategies.

*Stage Two.* The next stage of learning presents two main alternatives. Students may be allowed to construct new thinking strategies on their own, or they may be guided in this process. Previous studies have shown that even without any guidance, the percentage of valid inferences students make increases with time, indicating the acquisition of new thinking strategies. But these studies were conducted under research conditions and not in real classrooms where it seems unlikely that students will engage in one task for a period sufficiently long to induce change in reasoning strategies. Another study that compared independent with guided discovery learning in the context of our microworlds indicated that guidance contributes to improved performance as measured by the frequency of valid inferences. It was therefore decided to design guidance that will assist students in constructing the control of variables thinking strategy.

To begin with, it cannot be assumed that all students went through the same process as George, undermining their initial, invalid thinking strategies. So the teacher first directs all students to a set of experiences that are similar to the ones George had generated on his own. The teacher thus generates sets of uncontrolled experiments that naturally lead to contradicting conclusions. The contradictions produce hot debates among students about the correct conclusions and about the correct means for drawing these conclusions. Such a debate among peers serves to expose students' initial thinking strategies, to bring those strategies up in the public domain of the class, and to lay open conflicting views that may lead students towards cognitive dissonance with respect to suitable strategies for solving the problem.

With those conflicts in the background, the teacher then turns to help students construct new reasoning strategies. It is assumed that the same students who did not control variables in the context of the relatively complex photosynthesis problem, do have an intuitive understanding of variable control in everyday type of problems that consist of only two variables. The teacher then tells the following story:

John and Susan enter a room and Susan turns on the light. The light doesn't go on. John says: "Oh, it's the plug. It's not plugged in properly."

Susan says: "No, it's the bulb. The bulb is burnt." Then Susan changes the bulb and tightens the plug and the light goes on. So John says: "See—I was right. It was the plug." Susan says: "No, you were wrong. It was the bulb."

The teacher asks who is right, John or Susan? This question is followed by another class discussion. Then, depending on students' responses, the teacher might continue to ask one or more of these questions: "Can you tell for sure which of them is right? Why not? What would John and Susan have to do in order to know for sure who is right?"

Most junior high school students can explain it in this way: "We can't tell for sure who is right, because John and Susan changed two things at once." Our goal is to use this intuitive understanding as the basis for construction of a more general understanding of the rule of variable control. This is done by asking students whether they see any similarity between their thinking in the case of the photosynthesis problem and in the case of the story about the light bulbs. Although some students always answer that the similarity is that there are light bulbs in both stories, many detect the common thinking strategy and generate a response reflecting the need in both situations for changing only one factor at a time. Otherwise, students say, we can never know for sure what is the right answer. Thus, in their own words, students formulate their own rule of variable control. The teacher summarizes the discussion by introducing the formal term control of variables.

The aim of the subsequent, final stages of our learning sequence is to stabilize the use of the control of variables thinking strategy. Students are asked to bring examples of other incidents in which the same strategy should be used. Then several examples taken from their previous experiences with other scientific investigations as well as from other school topics and from everyday experiences are discussed. The goal of this stage is to prevent the welding of the strategy to the specific circumstances in which it was acquired and to enhance transfer.

*Stage Three.* Soon afterwards, students once again resume their independent investigations of the microworld. Consequently, they have multiple opportunities to practice the new rule and to stabilize its use.

During the second and third stages, special attention is given to metacognition. In the course of Stage Two, students are encouraged to reflect on their reasoning

strategies. The product of this process is the generalization or rule regarding the necessity of changing only one feature at a time. Whenever students fail to control variables during their independent investigation in the third stage, teachers once again direct them to reflect on the thinking strategies they have used, to compare them with the rule for the control of variables and thus to evaluate them.

While this learning sequence may be useful for most junior high school students, it is superfluous for some of them. Our assessment showed that approximately one tenth of the junior high school students tested had already mastered the thinking strategies for controlling variables before instruction took place. Those students certainly do not need three lessons on this subject, and are given instead an alternative assignment that consists of investigating interactions among variables in the photosynthesis microworld. The details of this assignment are beyond the scope of this chapter. It is, however, important to note that variability among students in initial reasoning abilities should be acknowledged and heterogeneous instruction may be designed to satisfy that variability. Another variation of the task designed for younger students consists of a similar but simplified microworld with only three variables. This version was used successfully in fourth, fifth and sixth grade.

Assessment. Students' progress was assessed through the comparison of a set of individual interviews conducted before learning took place, with a set conducted at the end of learning. An interviewer followed students' independent investigation of a problem, asking them to explain their inferences. Interviews were audiotaped and then transcribed and analyzed. Inferences were coded as either valid or invalid, according to the key of inference forms described by Kuhn, Schauble & Garcia-Mila. (1992). Altogether, sixteen eighth-grade students and seventeen ninth-grade students were interviewed.

In order to investigate transfer and retention, some students were interviewed twice more. For transfer, students were asked to investigate a new, logically equivalent problem in a different topic. In order to test retention of the acquired thinking skills across time, the eighth-grade students were interviewed again in the following school year, when they were in ninth grade, approximately five months after instruction had taken place.

The findings from this study showed that the rate of valid inferences increased from eleven percent in the early interviews to seventy-seven percent in the late interviews. Students were able to transfer their newly acquired reasoning strategies to a new problem taken from a new biological topic. They were also able to retain their newly acquired strategies across time, and to transfer them to yet another biological topic five months after instruction took place (Zohar, 1996).

## SUMMARY AND CONCLUSIONS

This chapter focuses on the observation that inquiry learning involves reasoning patterns that many students have not yet mastered. Consequently, when they engage in inquiry learning they often fail to understand some of the processes included in sound investigation. Rather than give up inquiry altogether or present algorithms that children must follow in a meaningless way, it is suggested to address reasoning difficulties as an explicit instructional goal and to teach accordingly.

Conversations with many science teachers reveal that they are well aware of their students' reasoning incapacities and occasionally address them in class. But they rarely think of them as distinct, explicit educational goals that must be addressed repeatedly and systematically. They usually plan lessons according to content objectives, and almost never plan lessons to address specific reasoning goals (Zohar, 1999). It may be naïve to expect that a single series of three lessons, such as the ones described here can indeed induce a significant change in the reasoning of students. It should be remembered that this series of lessons is part of a larger set of learning activities that science teachers who participate in the TSC project can incorporate constantly into the course of their instruction. Not all learning activities require three whole lessons. Some may consist of less time consuming means, such as a few questions, or a guided reflection upon reasoning processes that had taken place during inquiry. Addressing students' reasoning by pre-planned means, however, should become a routine part of instruction that is integrated constantly into science learning.

Conceptual change theories of learning, are based on a diagnosis of students' alternative concepts, followed by well-planned, structured, remedial instruction. Accordingly, sensitivity to students' reasoning difficulties should lead to a diagnosis of particular thinking problems that may then lead to instruction aimed to help students overcome those problems. Integrating such instruction into inquiry learning will help students to improve their understanding of scientific inquiry processes and thus to find them more meaningful.

## REFERENCES

- Adey, P. S., M. Shayer, and C. Yates 1989. *Thinking science: Student and teachers' materials for the CASE intervention*. London: Macmillan.
- Adey, P., and M. Shayer. 1993. An exploration of long-term far-transfer effects following an extended intervention program in the high school science curriculum. *Cognition and Instruction* 11: 1-29.
- Carey, S., R. Evans, M. Honda, E. Jay, and C. Unger. 1989. An experiment is when you try it and see if it works: A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*. Special Issue 11: 514-529.
- Carey, S., and C. Smith. 1993. On understanding the nature of scientific knowledge. *Educational Psychologist* 28: 235-251.
- Ennis, R.H. 1989. Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher* 18: 4-10.
- Friedler, Y., and P. Tamir. 1986. Teaching basic concepts of scientific research to high school students. *Journal of Biological Education* 20: 263-270.
- German, P. J., R. Aram, and G. Burke. 1996. Identifying patterns and relationships among the responses of seventh-grade students to the science process skill of designing experiments. *Journal of Research in Science Teaching* 33: 79-99.
- Inhelder, B., and P. Piaget. 1958. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Klahr, D., and K. Dunbar. 1988. Dual space search during scientific reasoning. *Cognitive Science* 12: 1-48.
- Klahr, D., and A.L. Fay. 1993. Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology* 25: 111-146.

- Kuhn, D., E. Amsel, and M. O'Loughlin. 1988. *The development of scientific thinking skills*. Orlando, FL: Academic Press.
- Kuhn, D. 1989. Children and adults as intuitive scientists. *Psychological Review* 96: 674-689.
- Kuhn, D., L. Schauble, and M. Garcia-Mila. 1992. Cross-domain development of scientific reasoning. *Cognition and Instruction* 9: 285-327.
- Lawson, A. E. 1995. *Science teaching and the development of thinking*. Belmont, CA: Wadsworth Publishing Co.
- Metz, K. E. 1995. Reassessment of developmental constraints on children's science instruction. *Review of Educational Research* 65: 93-127.
- Schuable, L. 1990. Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology* 49:31-57.
- Sodian, B., D. Zaitchik, and S. Carey. 1991. Young children's differentiation of hypothetical beliefs from evidence. *Child Development* 62: 753-766.
- Tamir, P., and V. N. Lunetta. 1978. An analysis of laboratory inquiries in the BSCS yellow version. *The American Biology Teacher* 40: 353-357.
- Tamir, P., R. Nussinovitz, and Y. Friedler. 1982. The design and use of practical tests assessment inventory. *Journal of Biological Education* 16: 45-50.
- Zohar, A., Y. Weinberger, and P. Tamir. 1994. The effect of the biology critical thinking project on the development of critical thinking. *Journal of Research in Science Teaching* 31:183-196.
- Zohar, A. 1994. Teaching a thinking strategy: Transfer across domains and self learning versus class-like setting. *Applied Cognitive Psychology* 8: 549-563.

Zohar, A. 1996. Transfer and retention of reasoning skills taught in biological contexts. *Research in Science and Technological Education* 14: 205-219.

Zohar, A. 1999. Teachers' metacognitive knowledge and the instruction of higher order thinking. *Teaching and Teacher Education* 15: 413-429.

Zohar, A., and F. Nemet. Submitted. Fostering students' argumentation skills through bioethical dilemmas in genetics.

# Teaching Science as Inquiry for Students with Disabilities

J. Randy McGinnis

---

## INTRODUCTION

Currently, 5.4 million students in American public schools are identified as having disabilities: impairments such as speech, hearing, motor or orthopedic, and visual difficulties, and conditions eligible for special education services such as learning disabled (LD), developmental delay or mental retardation (MR), autism, traumatic brain injury, and seriously emotionally disabled (EH). Data reported by the Department of Education in 1991 indicate that over half of students with documented disabilities receive instruction in regular education classes. This figure will increase as more school districts come into compliance with Public Law 101-476, the Individuals with Disabilities Education Act (IDEA) of 1997, by placing students with disabilities in the mandated “least restrictive environment.”

Teachers have customarily identified science classes as especially suited for inclusion of students with disabilities (Atwood & Oldham, 1985). They note the relevance of the content, the possibility of practical experience, and the opportunity for group learning with typical peers (Mastropieri et al., 1998). This perspective does not mean that most science teachers are sanguine about including

students with disabilities in their classrooms. Instead, as reported by Katherine Norman, Dana Caseau, and Greg Stefanich in 1998, both elementary and secondary school teachers identify as one of their primary concerns teaching students with special needs.

Recent standards for instruction as F. James Rutherford and Andrew Ahlgren (1989), the American Association for the Advancement of Science (1993), and the National Research Council (1996) present them put inquiry at the center. What could inquiry-based science instruction be for students with disabilities? For a beginning to an answer, I want to review four relevant fields of the literature: present-day portrayals of inquiry learning by writers addressing the teaching of science to students with disabilities; reasons for science instruction of that sort for such students; evidence that the approach is appropriate for them; and implications for teachers of science to students with disabilities.

#### PORTRAYALS OF INQUIRY LEARNING BY PROFESSIONALS CONCERNED WITH TEACHING STUDENTS WHO HAVE DISABILITIES

Among writers considering science instruction based in inquiry for students with disabilities are Thomas Scruggs et al., who in 1993 linked inquiry-based instruction with specific science curriculum projects such as the *Full Option Science System* (Lawrence Hall of Science, 1992) that emphasize student activities in small groups. Scruggs and Margo Mastropieri in 1994 identified such instruction with strategies endorsed by F. James Rutherford and Andrew Ahlgren in 1989: emphasis on concrete meaningful hands-on experiences in place of vocabulary acquisition and textbook learning. In 1995 Joyce Sasaki and Loretta Serna described such instruction in science as an inductive process that proceeds in a sequence referred to as “knowing and doing” (p. 14): establish what students already know, ask questions about what is to be observed, investigate, and obtain new knowledge. The appropriate role for the science teacher is that of facilitator.

In one of the most current studies reported, Mastropieri, Scruggs, and Butcher in a study published in 1997 identify instruction founded in inquiry as a type of inductive thinking that requires drawing general rules after observing a number of specific observations. They associate it with “constructivism” (p. 199), in which every learner constructs meaning individually in classrooms where the teacher engages in coaching and in promoting thinking activities. Bridget Dalton et al., (1997) describe this variety of science instruction as informed by the constructivist principle of social interaction. Support by the

teacher they view as critical to the process. This is so important a tenet that they name “supported inquiry science” their version of instruction by inquiry. Paramount is their idea of learning as conceptual change. A central assumption is that hands-on activities are not sufficient to change a learner’s thinking. The students should be asking questions, designing experiments, observing, predicting, manipulating materials, keeping records, and learning from mistakes (p. 670).

## REASONS FOR GIVING STUDENTS WITH DISABILITIES SCIENCE INSTRUCTION BASED IN INQUIRY

In the *National Science Education Standards*, issued in 1996 by the National Research Council (NRC), references to teaching students with disabilities unequivocally support including them in science classrooms that teach by inquiry and having them participate. A central principle guiding the development of the *Standards* is “Science for all students” (p. 19), defined as a principle of “equity and excellence” or fairness (p. 20). All students are also assumed to be included in “challenging science learning opportunities” (p. 20).

This equity principle is reflected in Teaching Standard B of the *Standards*. Teachers should recognize and respond to student diversity and encourage all students to participate fully in science learning (p. 32). Students with physical disabilities might require modified equipment; students with learning disabilities might need time to complete science activities (p. 37). The equity principle is also contained in Program Standard E: All students in the K-12 science program must have equitable access to opportunities to achieve the *Standards* (p. 221). Actions to promote this include bringing in students who have not customarily been encouraged to do science, among them students with disabilities, and making adaptations responsive to their needs (p. 221). This equity principle is further reflected in Assessment Standard D: Assessment practices must be fair (p. 85). Assessment tasks must be appropriately modified to accommodate the needs of students with physical disabilities (and) learning disabilities (p. 85). In particular, fairness requires students with disabilities to “demonstrate the full extent of their science knowledge and skills” (p. 86).

The ethical position taken in the *Standards* is in compliance with the constitutional reasons expressed in Public Law 101-336, the Americans with Disabilities Act of 1990, and in IDEA. IDEA is the most encompassing legislative victory by advocates for students with disabilities who have long fought for appropriate educational opportunities for all students. It is a telling repudiation of a mindset that got expression in 1903, when the Committee on Colonies for

Segregation of Defectives influenced the National Conference on Charities and corrections to campaign for the exclusion of students with disabilities in American schools (Gilhool, 1998). Of particular relevance in IDEA is the mandate to base on the content of the regular science curriculum all instruction for students with disabilities. This means that the Individualized Education Plans (IEPs) for students with disabilities must now describe curricular adaptations and accommodations based on the regular curriculum. So as the recommendation by the NRC to base science instruction on inquiry permeates the nation's science curricula, federal law increasingly supports inquiry for all students.

### EVIDENCE THAT INQUIRY IS APPROPRIATE FOR STUDENTS WITH DISABILITIES

The passage in 1975 of Public Law 94-142, the Education for the Handicapped Act intensified research on science instruction for students with disabilities (Bay et al., 1992). The provision mandated by the law to place handicapped students in the least restrictive environment meant suddenly that significantly more students with disabilities were in the general education science classroom. The chief concern was whether students with disabilities would benefit from their "selective placement...in one or more regular education classes" (Rogers, 1993, p. 1), and if so, which type of instruction is the most effective. Findings from these studies typically relate exclusively to students described with mild disabilities and concentrate on the impact of "mnemonic instruction, free study, direct questioning, and direct instruction" (Bay et al., 1992, p. 556).

In a comprehensive review published in 1992 of the literature from the 1950s until the early 1990s on science education for students with disabilities, Margo Mastropieri and Thomas Scruggs observe that from the early 1950s to the passage of P.L. 94-142 in 1975, research centered on "the effectiveness of developmentally oriented, hands-on curriculum to improve the content knowledge and cognitive functioning of students with disabilities" (Scruggs et al., 1993, p. 2). But, concurrently with the passage of P.L. 94-142 and the emergence of the back-to-basics movement, and extending until the 1980s, published studies on the developmental science activities for students with disabilities gradually stopped. Instead, increasing numbers of studies focused exclusively on basic skill acquisition by students classified as having learning disabilities.

Only in the 1990s, when the "commitment to educate each child, to the maximum extent appropriate, in the school and classroom he or she would otherwise attend" (Rogers, 1993, p. 1) became the norm, did research into the impact of

instruction by inquiry for students with disabilities get published again. But, as Thomas Scruggs and Margo Mastropieri reported in 1994, the research in the inclusion in the science classroom of students with disabilities had little to say of students with developmental, emotional, or behavioral as opposed to physical disabilities (p. 805). A review of recent studies looking into the effects of science instruction by inquiry on selective students with certain disabilities indicates that when the lessons are appropriately structured, students with learning disabilities benefit by the acquisition of knowledge; they also express greater satisfaction with hands-on science activities rather than with textbook activities (Scruggs et al., 1993). Students with visual, physical, auditory, or learning disabilities were evaluated by their teachers as successfully participating in all aspects of their elementary school science classes, including science activities, classroom discussion, and completion of adapted assignments (Scruggs & Mastropieri, 1994a). Students with various disabilities benefit from the use of technology to solve problems and to acquire and analyze data (Alcantra, 1996, as reported by Woodward & Rieth, 1997). For mastering concepts, students with and without learning disabilities profit from inquiry (Dalton et al., 1997). In courses centered in inquiry, students with learning disabilities, mental retardation, or emotional problems make academic gains comparable to those of their classroom peers and superior to the gains of most peers without disabilities who take courses based in textbooks (Mastropieri et al., 1998). In construction of scientific knowledge and in learning, remembering, and comprehending, inquiry aids more than direct provision of the same information (Scruggs & Mastropieri, 1994b). Students with learning disabilities may be able to participate in inquiry based on constructivist principles and benefit from it, but it is suggested that mentally retarded students may not benefit to a similar degree (Mastropieri, Scruggs, & Butcher, 1997).

Still needed to inform this research are many thoughtful case studies that examine specific disabilities in science learning based on inquiry. In their absence, there is no preponderance of evidence to indicate whether or not instruction by inquiry as guided by the *Standards* remains for students with disabilities a promising pedagogical initiative.

## IMPLICATIONS FOR INSTRUCTORS WHO TEACH SCIENCE TO STUDENTS WITH DISABILITIES

Students with disabilities are currently taught science in either self-contained or inclusion classrooms. Research based on survey methodology indicates that in both of these contexts teachers believe they are ill-prepared for the task (Holahan, MacFarland, & Piccollo, 1994; Norman, Caseu, & Stefanich, 1998). In both contexts, the *Standards* recommend instruction based in inquiry. The primary implication for teachers who teach science to students with disabilities is therefore to develop a vision for instruction by inquiry for all students. The emerging literature on science instruction by means of inquiry for students with disabilities provides an additional essential source of information from which to forge personal visions of inquiry. Here I present two examples.

In 1995 Sasaki and Serna reported on an inductive science program for middle school students with mild disabilities that they evaluated as effective. The program is titled Foundational Approach to Science Teaching I (FAST I). FAST I is described as a “hands-on, practical, inquiry approach” (p. 14) that teaches physical science, ecology, and relational topics. It features work by students in groups as well as experiments and public data charts. Students solve anomalies and are expected to make interpretations of the data, which includes making extrapolations and interpolations and drawing conclusions. A nonnegotiable portion of the instruction is clear direction on appropriate student participation. Recommendations are made for warnings and time-out procedures. Points are awarded for appropriate participation. The two teachers also recommend from experience requiring a notebook with a format; generation by students of presentations or hypothesis on what will be learned from the experiment; participation in experimental activities; oral presentations by the learning groups on the data they collected; elicitation of summary statements and group construction of conclusions. The authors assert that this approach develops critical thinking, enhances self-esteem, and furthers academic success.

For science classrooms employing inquiry and including students with disabilities and others without, a collaborative relationship with special educators is recommended (McGinnis & Nolet, 1995; Stefanich, 1994). A collaborative pairing makes for exchanging pedagogical and scientific knowledge. A model for bridging the space between science instruction and special education proposed by Victor Nolet and Gerald Tivnan as described by J. Randy McGinnis and Nolet in 1995 is directed toward achieving a specifically defined content. This model also includes problem-solving formats that emphasize concepts,

not rote memorization of facts. Expected of students are the intellectual operations of description and problem solving and the acquisition of an identified assemblage of facts, concepts, and principles.

For a lesson on fossil fuels, for example, a middle school science teacher, for example, might identify “acid preparation” as an essential concept. Inquiry would be guided by recommendations in the *Standards*: making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. The teacher will have established expectations for student assessment differing according to the IEP of the student. A student whose IEP targeted the intellectual descriptive operation would be expected minimally to summarize the defining attributes and provide illustrative examples of “acid rain.” A typical student in the same lesson might be expected to demonstrate intellectual evaluation in a task that requires taking a positive stance with a well-developed rationale toward the continued use of fossil fuels.

## CONCLUSION

This is a time of national reform in science education. The *Standards* published by the NRC urge that for all students instruction in science be by inquiry. While much is not settled on realizing such instruction for students with disabilities, intellectual ferment and practitioner initiative are in ample supply. The answer to what science instruction by inquiry could be for students with disabilities are multiple resplendent visions with more still to emerge during this time of remarkable opportunity.

## REFERENCES

- American Association for the Advancement of Science. 1993. *Benchmarks for science literacy*. New York: Oxford University Press.
- Atwood, R.K., and B. R. Oldham. 1985. Teachers' perceptions of mainstreaming in an inquiry oriented elementary science program. *Science Education* 69:619-624).

- Bay, M., J. R. Staver, T. Bryan, and J. B. Hale. 1992. Science instruction for the mildly handicapped: Direct instruction versus discovery teaching. *Journal of Research in Science Teaching* 29 (6), 555-570.
- Dalton, B., C. C. Morocco, T. Tivnan, and P.L.R. Mead. 1997. Supported inquiry science: Teaching for conceptual change in urban and suburban science classrooms. *Journal of Learning Disabilities* (November/December) 30(6):670-684.
- Gilhool, T.K. 1998. Advocating for our children: Will it every stop? *Down Syndrome News* (November) 22 (9):115-120.
- Holahan, C., J. McFarland, and B.A. Piccollo. 1994. Elementary school science for students with disabilities. *Remedial and Special education*. 15(2): 86-93.
- Lawrence Hall of Science. 1992. *Full option science system*. Chicago: Encyclopedia Britannica Educational Corp.
- Mastropieri, M.A., and T.E. Scruggs. 1992. Science for students with disabilities. *Review of Educational Research* 62(4): 377-412.
- Mastropieri, M.A., T.E. Scruggs, and K. Butcher. 1997. How effective is inquiry learning for students with mild disabilities? *The Journal of Special Education* 31 (2):199-211.
- Mastropieri, M.A., T.E. Scruggs, P. Mantziopoulos, A. Sturgeon, L. Goodwin, and S. Chung. 1998. A place where living things affect and depend on each other: Qualitative and quantitative outcomes associated with inclusive science teaching. *Science Education* 82:163-179.
- McGinnis, J.R., and V.W. Nolet. 1995. Diversity, the science classroom and inclusion: A collaborative model between the science teacher and the special educator. *Journal of Science for Persons with Disabilities* 3:31-35.
- National Research Council. 1996. *National science education standards*. Washington, DC: National Academy Press.

- Norman, K., D. Caseau, and G. Stefanich. 1998. Teaching students with disabilities in inclusive science classrooms: Survey results. *Science Education* 82:127-146.
- Rogers, J. 1993. The inclusion revolution. *Phi Delta Kappan* 11:1-5.
- Rutherford, F.J., and A. Ahlgren. 1989. *Science for all Americans*. New York: Oxford University Press.
- Sasaki, J., and L. Serna. 1995. FAST science: Teaching science to adolescents with mild disabilities. *Teaching Exceptional Children* 27 (4):14-16.
- Scruggs, T.E., and M.A. Mastropieri. 1994a. Successful mainstreaming in elementary science classes: A qualitative study of three reputational cases. *American Educational Research Journal* 31(4): 785-811.
- Scruggs, T.E., and M. A. Mastropieri. 1994b. The construction of scientific knowledge by students with mild disabilities. *The Journal of Special Education* 28: 307-321.
- Scruggs, T.E., M.A. Mastropieri, J. P. Bakken, and D. F. Brigham. 1993. Reading versus doing: The relative effects of textbook-based and inquiry-oriented approaches to science learning in special education classrooms. *The Journal of Special Education* 27(1):1-15.
- Stefanich, G. 1994. Science educators as active collaborators in meeting the educational needs of students with disabilities. *Journal of Science Teacher Education* 5(2): 56-65.
- U.S. Department of Education. 1991. *Thirteenth annual report to Congress on the implementation of the education of the handicapped act*. Washington, DC: US Government Printing Office.
- Woodward, J., and H. Rieth. 1997. A historical review of technology research in special education. *Review of Educational Research* 67(4):503-536.

## **Appropriate Practical Work for School Science— Making It Practical and Making It Science<sup>1</sup>**

Brian E. Woolnough

---

This chapter is partly a cautionary tale and in part exhortation. As a cautionary story, it looks back on the long tradition in the United Kingdom (UK) of practical work in science lessons and that too much of such practice trivialises science. But it is also an exhortation to continue personally motivating and fulfilling labor that introduces students to genuine, authentic science activity.

In this chapter I will be talking about the types of practical work that engages students, either individually or in small groups, with scientific or quasi-scientific apparatus. There is no single name for such activities. What is called an inquiry-based classroom in one country is called a practical laboratory in another; what some call hands-on experimentation others call inquiry learning and others still practical work, explorations or investigations. I hope that in the discussion below I will make clear what type of practical work or inquiry learning I am talking about, what its purposes are and what is effective.

I should say by way of introduction that I believe passionately in the importance of practical work in science teaching. In the UK much practical work is done in teaching science, many science teachers feel guilty if they are not teaching science through practice. Yet I think that much of the practical work that is done there in science lessons, and I would suggest in the science lessons of most other countries, is ineffective and detrimental to an appreciation and enjoyment of science.

Do we have problems with the practical work currently being done? If so, and I believe that we do, what is the cause? What type of practical work is appropriate and is it possible to do a type of science in schools that is authentic, introducing students to the way that many scientists actually work?

## PROBLEMS WITH THE CURRENT SITUATION

A recent leader in the 1995 British Council Newsletter said that much practical work is ineffective, unscientific, and a positive deterrent for many students to continue with their science. It is ineffective in helping students understand the concepts and theories of their science. It is unscientific in that it is quite unlike real scientific activity. And it is boring and time wasting for many students who find it unnecessary and unstimulating.

I agree with that: I wrote it! But I am not alone in my worries about the effectiveness of practical work in school science. Practical work as many schools employ it, observed Derek Hodson in 1992, “is ill-conceived, confused and unproductive. It provides little of real educational value. For many children, what goes on in the laboratory contributes little to their learning of science or to their learning about science. Nor does it engage them in doing science in any meaningful sense.”

An article of 1996 by Hodson describes the development of practical work since the 1960s as “three decades of confusion and distortion.” In the words of OFSTED (the UK government’s Office for Standards in Education) (1994), quoting Paul Black and Rosalind Driver and their group of leading science educators, “...there is a large body of evidence, both in this country and many others, that the understanding of the ideas of science which pupils develop within and at the end of their courses is alarmingly poor. . .”

## THE CAUSES OF THE PROBLEMS

Among the reasons why, I believe, much practical work is ineffective, is that many teachers and students appear to be uncertain about the aims and objectives for doing practical work in science lessons, and therefore turn inquiry into a mere exercise in busyness. We need to be clear about what the reason is for doing any particular practical task and alert the students to it. If we hope that a single practical work will fulfill a whole range of vaguely specified objectives at the same time we will probably be disappointed.

Another caveat is that two particular aims often get confused in a way that does serious damage to both. Is the aim of practical work to give an increased understanding of some theory or is it to develop some practical skills? Is the experiment being done for the sake of theory, to discover, to verify or to clarify it, or is it being done for the sake of developing ways of working like a scientist? If we try to do an experiment which fulfills both of those aims at the same time we will not succeed in either. To get the experiment to clarify the right theory we will have so to direct and constrain it that we kill any freedom that the students may have for developing their practical skills. The old Nuffield theory of guided discovery just did not work (Stevens, 1978; Hodson, 1996).

There has been confusion here for a long time. A recent research investigation that I carried out with teachers and students in Oxford asked, among other things, for teachers to rank a series of aims for practical work in order of importance and also to rate how frequently they did different types. The most important objectives in their judgement related to developing practical skills; using practice to discover or clarify theory they rated very low. And yet the type of practice which teachers said that they used the most were “structured practical linked to theory.” This finding confirms earlier research (Kerr, 1964; Thompson, 1975; Beatty et al., 1982). Another recent survey (Watson, 1997) asked not only the teachers but the students what they thought the purpose of doing practical work in science was. The teachers split their aims almost evenly between the development of theoretical understanding and the crafting of practical skills. But all the students believed that the only reason for doing practical work was to increase their theoretical knowledge and understanding! We must help our students, as well as our teachers, to distinguish between learning to understand scientific knowledge and learning to do scientific activity.

A further cause of confusion in the use, or misuse, of practical work is the information overload which bombards students, of which we are often unaware. As teachers, we know why we are doing a practical inquiry and see clearly the theory underlying the experiment. But students will have to sort out the apparatus, remember how it works, select the relevant data from the experiment, ignore the irrelevant, take measurements to the appropriate degree of accuracy, handle those inevitably imprecise and messy data, analyse into a meaningful conclusion the information gathered and then, after so much information has come in, try to remember what the experiment was all about in the first place—what question were they trying to answer. Often the underlying principle—say, Newton’s laws of motion—is elegantly simple but by the time

we have wrapped it up with carts, springs, and dotted ticker tape we have produced a highly complicated and distracting mass of clutter which hides it.

Another reason why much practical work is ineffective is that the preconceptions the students bring with them to the experiment determine what they will see. Just as we, the teachers, see the correct theory shouting at us through the experiment, so the students will see what they think is correct. In an electric circuit with two light bulbs in series, we would anticipate that the bulbs are lit equally brightly, and if one is slightly brighter than the other we conclude that the bulbs are not identical. But if we, as students, expected the current to be used up going round a circuit, we would expect one bulb to be brighter than the other and would convince ourselves that it was so. The POE (Predict, Observe, Explain) strategy (Gunstone, 1991) is useful here, forcing students first to predict what will happen, then observe and then explain the observation. This compels the students to engage with the experiment. As Joan Solomon said in 1980 when analysing why some of her experiments were successful for her science students and others were not, "Imaginative insight was not a sequel to successful experiments. On the contrary, it was an essential prerequisite."

A final and perhaps the most important reason why so many students gain so little from the practical work is that they come to the experiments without any intellectual curiosity, purpose or motivation. They come casually to the lesson, do what the teacher tells them to do and go away. I believe that almost any experiment can be effective if the student is genuinely motivated to find out what is going on and is determined to succeed. This entails ensuring that the students have a sense of ownership, that it is their experiment and not the teacher's.

## A RATIONALE FOR PRACTICAL WORK IN SCIENCE

So far I have been negative about much practical work. Let me now be positive and suggest what I believe is a rationale for practical work which is thoroughly constructive and also involves the students in doing real science. Fifteen years ago, when sorting out my own thinking on practical work in science, I established three clear principles. I wanted greater clarity for the reasons for doing practical work, and a matching of project to aim. I wanted to separate theory from practice, employing practice not to discover or elucidate theory, but for its own sake. And I wanted a threefold rationale for practical tasks: experiences to develop a feel for the phenomena being studied; exercises to develop practical skills; and investigations to develop experience and expertise at working like a problem-solving scientist, which involves planning, performing, interpreting,

and communicating. I still think that was a useful starting place but would now want to develop the framework a little more.

My rationale for practical work in school science would now be based on this framework:

- ▶ exercises to help develop practical skills;
- ▶ experiences to give students a feel for the phenomena;
- ▶ scientific investigations that include problem solving to gain experience of being a problem-solving scientist and often using technological, open-ended investigations and hypothesis testing to develop the way of working as a hypothesis tester and often using pure science and closed investigations;
- ▶ demonstrations to develop a theoretical argument and to arouse interest and to make an impact;
- ▶ recipe experiments to keep pupils occupied and to personalize some theory with POE.

There are certain scientific skills, sometimes called the processes of science, which need to be learnt, such as taking measurements, using scientific equipment, analysing data, tabulating and interpreting graphs, and these need to be developed with appropriate **exercises**. Sometimes these exercises will form an integral part of a larger scientific experiment, and be developed as occasion arises, but it should be made clear to the student what the aim actually is. Some educators talk about process skills, such as planning, hypothesising, observing, analysing, and interpreting, and suggest that these can be taught and assessed separately, out of context, and then put together to form the whole process through which a mature scientist works. But as Millar and Driver so clearly argued in 1987 and Millar in 1991, such processes are dependent on context. Being good at observing scientifically is not the same as being observant. The broad process skills of science can be properly developed only in the context of an authentic scientific activity

**Experiences** are invaluable for gaining a feel for the phenomena being studied, as opposed to developing the underlying theory. Such experiences will often be very simple, short and straightforward; stretching an elastic band to get a sense of elasticity; burning magnesium ribbon to get a sense of combustion; dissecting a plant to get a sense of the size, the shape, and the texture of the component parts; moving your finger around and above the lighted flame of a candle to get a sense of convection. These experiences can then be the basis for thinking about the underlying theory—but that will have to be done theoretically, by question and answers, discussion and exposition.

**Investigations** come in different forms. They may be closed or open-ended, involving finding what factors affect the strength of an electromagnet, which material makes the best sole for a shoe or how to make a parachute which lowers an egg from the top of a science block roof onto the ground in the quickest time without breaking the egg. They may take a short time or be spread over a week or so. They may fit into the normal science curriculum or be extracurricular. They may be based in a lab or the environment. But all should start with a question or a problem and get the students to make their own plans for tackling it. This very act of having responsibility for the planning of the investigation gives the students ownership of it and a determination to make it work. They will then perform the investigation, analysing and displaying their results as appropriate, and then evaluate and possibly modify their experiment. There is no single way of doing a scientific investigation. It involves the affective as well as the cognitive; it demands creativity, guesses, hunches, experience, a whole range of tacit knowledge and perseverance—and all these can be developed through investigations. Creativity cannot be taught, but it can be encouraged. Scientific investigations encourage creativity in a way little else in science teaching does. I have written more about scientific investigations elsewhere (Woolnough, 1994) and the ways that these can be incorporated in school science, both within the normal science syllabus and as extracurricular science activities. Suffice to say here that I believe that the opportunities for students to do their own investigations as a science project do more to transform a school's science course, and the students' enjoyment of it, than any other single thing.

A particular form of investigation which we, through the English National Curriculum for Science, have had experience of is a form of hypothesis testing. The original version in 1991 of the practical Attainment Target, Sc1, was called Scientific Investigation and took a topic from the physics, chemistry or biology programme of study and set up a rather closed investigation in which the student had first to set up an hypothesis as to what was likely to happen, to do the investigation and then to evaluate how far the hypothesis was proved correct. This led to some closed verification experiments—what factors, for example, affect the rate of reaction of metals and acids, what factors affect the strength of an electromagnet—which tightly prescribed what would happen and prevented genuine scientific activity. I have little time for such restrictive investigations, which seem to me to represent a very poor model of what scientists really do. The latest version of our national curriculum, produced in 1995, allows a rather wider interpretation of scientific activity which includes both investigations and experimentation based on the processes of planning, obtaining evidence, analysing evidence

and drawing conclusions and evaluating evidence. This encourages a far wider range of practical activity, developing the skills and processes, and the conceptual understanding, in the most appropriate way.

**Demonstrations**, which I believe are more popular in the science classes of other countries than in England, are an excellent way of developing a theoretical argument in which the teacher can establish the structure of the relationships in a way which would not be possible for students left to sort out the ideas for themselves. Class demonstrations, often with student involvement, are an ideal way for the teacher to link together the theoretical and the practical,

Another class of experiment which I would refer to as **recipe experiments** has students merely follow instructions. I have little enthusiasm for these, which so often put students to just doing what they are told without understanding, insight or ownership. If such experiments are to be useful then they need to be focused. Somehow the students need to be motivated and appreciate what the experiment is about. Utilising the POE technique to force the students to think and become committed to their results, will perhaps give them a clearer impression of the significance of the experiment.

#### AUTHENTIC SCIENCE, PERSONAL KNOWLEDGE, AND SCHOOL SCIENCE

Much of scientific knowledge comes in the form of public knowledge. Knowledge that is public is written down in books and journals, is described in syllabuses, and provides a basis for public examinations. Public knowledge is explicit and makes possible a community of effort.

But there is another type of knowledge which is personal, gained through private experience and practice, and less easy to define. It relates both to understanding aspects of the physical world and also to the procedures required to tackle problems in it. It is tacit, and involves our senses as well as our intellect, our emotional commitment along with our intelligence. In Polanyi's words in his seminal book on personal knowledge published in 1958, "We know more than we can tell." We assert that a piece of music is by Bach even though we have never heard it before, we know how to ride a bicycle even though we cannot explain its stability and why we don't fall off, we recognise a suitable substance and a sensible course of action in tackling a scientific problem even though we cannot always explain why. Guy Claxton argued in 1997 that "we should trust our unconscious to do the thinking for us." Scientists in their professional life rely on their personal knowledge, their intuition, and their creativity as much as their

public knowledge, if not more. Indeed it is only as the public knowledge has become personalised that it can be used. But personal knowledge is more than just personalised public knowledge: it is deeper, more sensual, inarticulate, and yet most useful.

School science in most countries is increasingly dominated by accountability and assessment and thus the form of science which is taught is predominantly public knowledge. Personal knowledge, because the individual cannot make it reliably explicit, gains little credit or status. But I would argue that unless school science teaching allows some place for the development and expression of personal knowledge in science, we will be misleading students about the authentic nature of scientific activity. I believe that we need also to stress the difference between the two types of knowledge and stress that both have value and importance. In 1971 J. Ravetz spoke of science as “a craft activity, depending on a personal knowledge of particular things and a subtle judgement of their properties.” Learning in science W. M. Roth described in 1995 as “an enculturation into scientific practices.” Millar asserted in 1996 that “learning science is a process of coming to see phenomena and events through a particular set of spectacles, and the intention of science education is to bring the learners inside the particular, and peculiar, view of things which scientists, by and large, share.”

Including authentic science activity of a practical problem-solving kind enables students to develop and use their personal knowledge through experience. It allows them to experience doing authentic science and thus partake in one of our principal cultural activities. It provides students with skills and attitudes, the personal knowledge, that are useful to employers. And motivates them towards science and hence increases their propensity to learn public knowledge of science too.

Employers, as well as society at large, want school leavers to have personal life skills and self-confidence enabling them to work as autonomous, self-motivated citizens. A survey by M.C. Coles in 1997 asking employers in industry, most of it based in science, what they wanted in their recruits, found a very large consensus. He found that they wanted:

- Commitment and interest (the most important criteria),
- Core skill capabilities such as communication, numeracy, and IT (Information Technology) capability,
- Personal effectiveness, relationships and team work,
- Self reliance, resourcefulness,
- Initiative, creativity,

- ▶ The skill of analysis,
- ▶ Good general knowledge including understanding of the world of work,
- ▶ Professional integrity.

Employers were not just interested in their applicants' having a large amount of scientific knowledge or understanding, or specific scientific skills, though a basic amount of both was required. They asked for broad personal core skills and attitudes. These are exactly the type of skills and attitudes that are developed when students do personal investigations in their science practical work. A recent evaluation of the engineering education schemes being done in schools in Scotland demonstrated that the student research projects within them greatly enhanced the skills of communication, interpersonal relations, and problem solving of the students involved.

### CREST,<sup>2</sup> AN EXAMPLE OF REAL PRACTICAL WORK

One of the most impressive and successful initiatives in the United Kingdom encouraging students to become involved in genuine scientific and technological activity has been the national CREST award scheme (CREST = CREativity in Science and Technology). CREST's aim is to stimulate, encourage, and excite young people about science, technology, and engineering through project work centered in tasks. It stimulates and supports projects connected to industry or community that draw on students' creativity, perseverance, and application of scientific and technological challenges. The projects cover a wide range of topics. Typical efforts might study the pollution of a local river, design and make an auto pilot for a sailing boat, or investigate the efficiency of alternative energy supplies. CREST gives bronze, silver, and gold awards. I have recently had the task of evaluating CREST after ten years of its existence and it quickly became apparent, through the response of both the teachers who organised it and the students who took part in it, that it was immensely impressive and highly successful in fulfilling its aims. Though CREST is an optional activity, involving only about 25,000 students—it is done by a minority of students in a minority of schools—those schools and students who are involved in it gain an enormous amount both in improving students' attitudes towards science and in developing their communication, interpersonal, and problem-solving skills.

My earlier researches in FASSIPES (a research project on the Factors Affecting Schools' Success In Producing Engineers and Scientists)(Woolnough, 1994b) showed that one of the factors which switched many students onto science

and technology was involvement with student research projects, and such extracurricular activities in science. One way of doing this was through CREST. It and similar ways of doing projects in and out of the curriculum are described in the book *Effective Science Teaching* (Woolnough, 1994a). Others have written of the effectiveness of such investigational projects in different countries. R. Tytler's account in 1992 of the Australian Science Talent Searches is especially convincing. In the States there is a similar Science Talent Search for many years sponsored by Westinghouse and now by Intel; South Africa has Young Scientists; Scotland has its Young Engineers competitions. Many countries have science clubs, science workshops, and "great egg races." I would rate this type of practical work as being the best way of developing all the skills, educational and vocational, that could be beneficial to responsible citizens, whether or not they were to enter scientific and technological industry—and doubly so if they were!

#### KNOWLEDGE, SKILLS OR ATTITUDES?

Much of what I have said about effective practical work has centred on the affective domain, on the principle of motivation. Motivation is important purely in personal terms, in developing and expressing a sense of self-worth. But motivation is also vital for gaining any understanding and appreciation of science, both as a body of knowledge and as a way of working. It matters not what students know, understand, and can do; what is important is what they *want* to do. Motivation, challenge, ownership, success, and self-confidence: These are much more vital outcomes than the acquisition of specific knowledge or skills. Many are sceptical of transferable skills which may be useful in broader aspects of life. Derek Hodson shared that scepticism but said in 1988

What may be transferable are certain attitudes and feelings of self worth...successful experience in one experiment may make children more determined and more interested in performing another experiment. The confidence arising from successfully designing an experiment might be a factor in helping children to stay on task long enough to design a new experiment successfully. (p. 260)

The Nuffield A level Physics curriculum of 1985 has a practical investigation as part of its course, and this is often the highlight of students' academic work. Students chose their own investigation, which typically lasts twenty hours. It could involve an investigation of water flows from pipes, an investigation of the acceleration of athletes, or the effect of stirring on the viscosity of non-drip paint,

or the factors affecting the lift on an airfoil. Such practical investigations are not always easy, but almost without exception they enable students to produce work of a higher quality than in any other aspect of their school work. For when motivated, students are without limits in what they can achieve in their scientific and technological projects (Woolnough, 1997).

## ENDNOTES

1. In the United Kingdom (UK) we refer as practical work to the type of activity done in laboratories based in inquiry. As much of this chapter will be a reflection of the UK experience (with the implications to be drawn for other countries left to the reader!), the term “practical work” will be used.
2. CREST stands for CREativity in Science and Technology. Details about the program are available from the CREST National Centre, 1 Giltspur Street, London EC1A 9DD.

## REFERENCES

- Beatty, J.W., and B.E. Woolnough. 1982. Practical work in 11-13 science: The context, type and aims of current practice. *British Educational Research Journal* 8(1): 23-30
- Claxton, G. 1997. *Hare brain, tortoise mind—why intelligence increases when you think less*. London: Fourth Estate.
- Coles, M.C. 1997. Science education—vocational and general approaches. *School Science Review* 79 (286): 27-32.
- Gunstone, R.F. 1991. Reconstructing theory from practical experience. In *Practical science*, edited by B. Woolnough, 67-77. Buckingham: The Open University Press.
- Hodson, D. 1988. Experiments in science and science teaching. *Educational Philosophy and Theory* 20: 253-66.

- Hodson, D. 1992. Redefining and reorienting practical work in school science. *School Science Review* 73(264): 65-78.
- Hodson, D. 1996. Laboratory work as scientific method: Three decades of confusion and distortion. *Journal of Curriculum Studies* 28(2): 115-135.
- Kerr, J.F. 1964. *Practical work in school science*. Leicester: Leicester University Press.
- Millar, R., and R. Driver. 1987 Beyond processes. *Studies in Science Education* 14: 33-62.
- Millar, R. 1991. A means to an end: the role of processes in science education. In *Practical Science* edited by B. Woolnough. Buckingham: The Open University Press.
- Millar, R. 1996. In pursuit of authenticity. *Studies in Science Education* 27: 149-165.
- Nuffield. 1985. *Advanced level physics. Examinations and investigations*. Rev. London: Longman.
- Office for Standards in Education. 1994. *Science and mathematics in schools: A review*. London: Her Majesty's Stationery Office.
- Polanyi, M. 1958. *Personal knowledge*. London: Routledge and Kegan Paul.
- Ravetz, J. 1971. *Scientific knowledge and its social problems*. New York: Oxford University Press
- Roth, W.M. 1995. *Authentic school science: Knowing and learning in open-inquiry science laboratories*. The Netherlands: Kluwer Academic Publishers.
- Solomon, J. 1980. *Teaching children in the laboratory*. London: Croom Helm.

Stevens, P. 1978 On the Nuffield philosophy of science. *Journal of Philosophy of Education* 12: 99-111.

Thompson, J. (Ed.) 1975. *Practical work in sixth form science*. Oxford: Oxford University Department of Educational Studies.

Tytler, R. 1992. Independent research projects in school science: Case studies of autonomous behaviour. *International Journal of Science Education* 14(4): 393-412.

Watson, J. 1997. Unpublished internal paper. London: Kings College.

Woolnough, B.E. 1994a. *Effective science teaching*. Milton Keynes: Open University Press

Woolnough, B.E. 1994b. Factors affecting students' choice of science and engineering. *International Journal of Science Education* 16(6): 659-676.

Woolnough, B.E. 1994c. Why students choose physics, or reject it. *Physics Education* 29:368-374.

Woolnough, B.E. 1995. Switching students onto science. *British Council Science Education Newsletter*. London: British Council.

Woolnough, B.E. 1997. Motivating students or teaching pure science? *School Science Review* 78(285): 67-72.

## Assessing Inquiry

Audrey B. Champagne, Vicky L. Kouba, and Marlene Hurley

---

Across the nation, science teachers are being called upon to use inquiry as a method of teaching and to enhance their students' ability to inquire. The American Association for the Advancement of Science (AAAS) and the National Research Council (NRC) have made inquiry an essential goal of the contemporary standards-based reform movement in science education. The AAAS *Benchmarks for Science Literacy*, published in 1993, and the NRC *National Science Education Standards*, which appeared in 1996, advocate the engagement of students in science inquiry so that they can develop both the ability to inquire and understanding of scientific principles and theories. The *Benchmarks* and *Standards* also contain descriptions of knowledge about inquiry and abilities to conduct inquiries that all students should gain as a result of their school science experiences. While the *Benchmarks* and the *Standards* differ somewhat in the emphasis placed on inquiry as a goal of science education,<sup>1</sup> each makes inquiry central to students' opportunity to learn<sup>2</sup> and the primary goal of contemporary science education reform. The *Benchmarks* and *Standards* serve as important guides to teachers who have the responsibility to meet national, state, or local science education standards.

This chapter describes assessment practices appropriate for science inquiry. It also describes how teachers can use assessment to increase their students'

opportunity to learn to inquire. These practices have their foundation in the principles of sound assessment practices contained in the assessment standards in the *National Science Education Standards*.

## INTRODUCTION

“Assessment,” “testing,” “evaluation,” and “grading” are often used synonymously. The meaning of assessment as we use it in this chapter is much broader. For our purposes, assessment is the process of collecting information for the purpose of making educational decisions. Administrators, government officials, teachers, and other educators gather information about what happens in science classrooms and what students learn as a result of their classroom experiences. Administrators and government officials primarily use very general information about student achievement to make judgements about the quality of teaching and schools, to allocate resources, and to make educational policy. Teachers gather more detailed and reliable information about student achievement and use it to plan courses and lessons, to monitor student progress, to inform students about the quality of their work, and to inform parents about their children’s work. The information collected by teachers has the greatest potential to increase the students’ opportunity to learn science and, in turn, to improve student achievement.

Information collected outside the classroom has profound implications for the lives of science teachers and their students. This information is collected in ways different from those employed by teachers within the classroom and is used in ways that vary from classroom purposes. Certain information is collected in controlled, rigorous ways and provides data that the district uses to communicate its educational status to federal and state officials and to the members of its community. Student performance on standardized science tests and per capita expenditures on science equipment are examples of data gathered by districts that are used for these accounting purposes. Federal and state agencies use information of this kind to allocate resources to districts. States use these kinds of data to sanction districts that are not performing up to standards. Community members use this information when voting for school tax bonds and school board members.

Because information collected outside the classroom has both direct and indirect influences on the school science program and the classroom practice of science education, science teachers must monitor and be involved in the planning and execution of these forms of assessment. But the focus of this chapter is on the information district personnel and teachers collect and use to make district-wide and classroom decisions; especially those related to students’ understanding

of scientific inquiry, their ability to inquire in scientific ways, and the opportunity schools afford students to develop these understandings and abilities.

## DEFINING INQUIRY

What evidence will teachers, administrators, and taxpayers accept that the students have learned to inquire and how will that evidence be collected? Defining expectations for performance and the conditions under which it is observed are parts of the assessment process. The nature of the evidence, in turn, depends on the district's expectations for the performance of students who have met successfully the district's standards. The expectations will reflect the district's definition of inquiry.

The *Benchmarks* and *Standards* provide a starting point for definition.<sup>4</sup> What is implied in both, but not made explicit, is that the practices of scientists and the scientific community serve as standards for the levels of inquiry expected of adults literate in science<sup>5</sup> and of students at various stages in the development toward adult science literacy.

The *Standards*, for instance, define as a component of science literacy the ability to ask questions. The rationale for this standard is that scientists pose questions and design inquiries, experiments, and investigations to answer them. While the practice of scientists is the implied standard, neither the attributes of the scientists' questions that makes them appropriate to scientific inquiry nor the scientific community's standards for question quality are defined explicitly in the *Standards*. The expectations for the levels of inquiry abilities and associated habits of mind to be attained by elementary, middle, and high school students identified in the *Benchmarks* and the *Standards* are not detailed enough to serve as the basis for the development of an inquiry assessment plan. Consequently, teachers must develop their own expectations building from these documents.

The *Benchmarks* and *Standards* employ the terms "scientific inquiry," "inquiry," "investigation," and "experimentation" with little indication of differences, if any, among them or indication of how they are related. More precise definition of these terms is a necessary condition for the development of strategies to assess them. In the absence of detailed definitions in the *Benchmarks* and *Standards*, we propose definitions on which we will base our discussion of strategies. Teachers may not agree with the distinctions we make among the different forms. That is all well and good. The essential principle is that the assessment strategy be congruent with the teachers' view of the nature of inquiry that they aim to develop in their students.

### Forms of Inquiry

Table 1 presents a framework for inquiry with summary descriptions of two major forms of inquiry. The feature that distinguishes these two forms is the practitioners. Scientific inquiry is practiced by natural scientists. Science-related inquiry is practiced by adults and students who are science literate. Table 2 differentiates three forms of science-related inquiry distinguished primarily by the purposes the form of inquiry serves. In turn, the purpose determines the design of the inquiry and the kinds of data collected.

**TABLE 1. TWO MAJOR FORMS OF INQUIRY**

FORM	PURPOSES	PRACTITIONERS
SCIENTIFIC INQUIRY Inquiry as Practiced by Natural Scientists	Understand the natural world; Formulate, test, and apply scientific theories (as in making designer drugs)	Research Scientists
SCIENCE-RELATED INQUIRY Inquiry as Practiced by Science Literate Adults and Students	Obtain scientific information necessary to make reasoned decisions;  Understand the natural world	Science Literate Adults  K-12 Science Students

**TABLE 2. THREE FORMS OF SCIENCE-RELATED INQUIRY**

FORM	PURPOSES	PRACTITIONERS
Archival Investigation	Collect information from print, electronic and other archival sources on which to base personal, social, civic decisions  Evaluate claims	Adults literate in science  K-12 science students
Experimentation	Testing of theory	Secondary science students
Laboratory Investigation	Develop understanding of the natural world Develop understanding of inquiry Develop abilities of inquiry Collect empirical data on which to base personal, social, civic decisions Evaluate claims	K-12 science students

### Scientific Inquiry

Scientific inquiry has as its primary purposes understanding the natural world by formulation and testing of theory and applying that understanding to practical problems such as the synthesis and testing of new drugs. The standards of practice of scientific inquiry are defined by the community of scientists, which specifies, among other things, the characteristics of empirical studies, methods of data collection, rules of evidence, and proper forms of reporting the results of experiments. Practices in the community of scientists are highly formalized and define a culture of science that includes a system of values, modes of

behavior appropriate to scientific discussions, and characteristic vocabulary and forms of expression.

### Science-Related Inquiry

The three forms of science-related inquiry are archival investigation, experimentation, and laboratory investigation. Archival investigation is practiced by adults who are literate in science and science students. Science literate adults inquire by gathering scientific information from paper, electronic, and other archival sources to make day-to-day personal, social, and economic decisions. Students also engage in archival investigation both in their personal lives and in their science programs.

In science class, students engage in science-related inquiry using empirical methods—employing observation, data collection, and the use of scientific apparatus and measuring instruments. Experimentation and laboratory-based investigations are two types of science-related inquiry that use empirical methods.

Experimentation is an inquiry for the purpose of testing a hypothesis that derives from a scientific theory.<sup>6</sup> An example is the test of a hypothesis that a mixture of carbon dioxide and natural gas behaves as an ideal gas. The hypothesis in the example comes from a theory of the behavior of gases. The experiment will test the behavior of the mixture of gases under certain conditions of temperature and pressure. If under the conditions tested the mixture behaves as an ideal gas, the hypothesis is supported. If, however, the gas mixture does not behave as an ideal gas, then the hypothesis is not supported. In neither case does the experiment prove or disprove the theory. It only demonstrates that under the conditions tested the results of the experiment support or do not support the theory. An instance of an investigation, also conducted in a laboratory, is a test to determine which of two brands of paper towels is more absorbent for the purpose of deciding which is the better buy. The question it is designed to answer does not relate to a theory of absorption or capillary action. The answer does not provide any insights into the nature of the natural world. Investigations require only that the test be fair and unbiased, while experiments must meet the rigid methodological requirements of the scientific community.

Before teachers can plan how to assess inquiry, they must decide which form of inquiry they expect their students to learn. Is the goal to develop graduates who will use the methods of scientific inquiry to develop new scientific knowledge or graduates who will have the inquiry abilities related to science that contribute to a satisfying and productive life? The answer to this question defines

the form of inquiry a science program seeks to engender and consequently the nature of the evidence teachers will collect as proof that the goal has been met. The choice will also influence forms of inquiry that will be emphasized in the program's assessment plan.

If the goal is to develop a set of inquiry abilities, an assessment task for the members of a science class might be, as part of making a decision about selection of vitamin supplements, to plan an archival investigation to learn what vitamin supplements scientists and physicians recommend. If the goal is to cultivate the ability to conduct laboratory investigations, the task for a middle school student might be to plan and conduct an investigation to test the absorbency claim of a paper towel advertisement and, on the basis of the results of the test, decide which brand of paper towels to purchase. If developing new knowledge is the goal, the task for a high school student might be to design and conduct an experiment to determine the degree to which a mixture of gases behaves as an ideal gas. All are examples of inquiry but, if we are assessing the ability of a student to inquire by use of these examples, each will have its own performance standards. These can be effectively generated through an examination of each phase of any inquiry process.

## PHASES OF INQUIRY

All forms of inquiry proceed in phases that we define as precursor, planning, implementation, and closure or extension. Within each of these phases, the student engages in two distinct aspects of inquiry: generation and evaluation. The student generates questions or hypotheses, plans for an investigation or experiment, and reports on the investigation or experiment. The student also evaluates the questions or hypotheses, inquiry plans, and reports of other students.

### General Description of Inquiry Phases

In the precursor phase, experiences in the natural world, in reading, in interactions with others, or in response to personal needs make students aware of something related to science about which they are motivated to learn more. Through interactions with others, the student refines the question or hypothesis and enters the planning phase. There the student collects information, refines the hypothesis or question, and selects an appropriate method for the experiment or investigation. The student also presents the plan to others, criticizes the plans of other students, and makes refinements. Some preliminary or pilot laboratory work may be done

as a part of the planning process. When a satisfactory plan is in place, the inquiry is conducted; information and data are collected and analyzed. It is possible during this implementation phase that problems with the plan may arise that require modifications. For instance, the need for additional data may become evident so that new data are collected and analyzed. During the final phase of the inquiry, the conclusions are evaluated to determine whether the inquiry has reached a satisfactory conclusion or further inquiry is required.

Successful engagement in inquiry requires knowledge of the nature of scientific inquiry, a variety of abilities including physical skills, and certain habits of mind. The inquirer must be able to pose questions and hypotheses amenable to scientific investigation and experimentation, to criticize plans for scientific investigations and experiments, to write a report about an inquiry, and to identify reliable sources of scientific information. These abilities, in turn, require information and the mental capabilities to process that information: that is, to reason scientifically.

### Assessment Investigations and Phases

Our discussion of assessment strategies within the four phases of inquiry uses as an example an investigation conducted primarily in the laboratory but gathering some information outside it. The purpose is to gain information on which to base a decision. Investigation based in the laboratory is a form of inquiry that many districts have identified as the inquiry goal for their science programs. Investigations within the laboratory of social or other problems go a long way toward meeting the Inquiry Content Standard of the *National Science Education Standards* and the inquiry goal in the *Benchmarks*, which states:

Before graduating from high school, students working individually or in teams should design and carry out at least one major investigation. They should frame the question, design the approach, estimate the time and costs involved, calibrate the instruments, conduct trial runs, write a report and finally, respond to criticism. (AAAS, 1993, p. 9)

While much of the knowledge and many of the abilities required for the investigations founded in the laboratory are the same as for experimentation, the two forms of inquiry differ in purposes (as shown previously in Table 2), in methodology (for instance, in the control of variables), and in the language for reporting results. Consequently, the discussion about investigation carried out in the laboratory cannot be applied to experiments without modification.

To assess the ability to inquire, teachers might ask students to design an investigation testing the claims made in breakfast cereal commercials about the vitamin and mineral content of the products. The results of the investigation could be used to decide which cereal to purchase. Do the products accurately report the USDA minimum daily requirements for each of the nutrients? This is a question that requires information gathering. Do the products actually contain the amount of each of the nutrients reported on the label? This might be answered empirically. To do so, the student must learn about how the nutrients are assayed and determine whether the chemicals and equipment necessary for the assays are available. What for each of the products is the ratio of nutrient to serving? Is the serving determined by volume or by weight? What is the ratio of serving to cost? What about taste appeal? Will the decision about which product to buy be based only on scientific information and cost? Not likely.

Assessment of a complete investigation of this type is intensive in resources. But if the ability to inquire is a central goal of a school science program, then students should have the opportunity to learn to conduct such investigations and be assessed on their ability to do so. Giving students the opportunity requires considerable time and equipment in laboratories as well as computers. If inquiry is to be a goal of a science program, these are considerations that should be taken into account. Once the decision is made, then districts must be ready to provide students with both the time and the equipment necessary to develop the ability, along with the opportunity to demonstrate their attainment of the ability. An alternative inquiry goal requires only that students know about scientific inquiry. That goal is less intensive in resources and the assessment strategies demanding only techniques for gathering information by paper and pencil.

If a district program is in compliance with the recommendations contained in the *Benchmarks* and *Standards*, much of the twelfth-grade science course for all students will be doing an extended investigation. Data collected as students engage in the investigation may be used to grade the students or to evaluate the effectiveness of the K-12 science program. Or as suggested in the *Benchmarks*, the purpose of engaging students in a complete investigation is to provide the students with the experience of putting together into a coherent whole the various elements of inquiry to which they have been exposed throughout the K-12 program. Either way, doing a complete investigation should be viewed as an occasion for learning. As some assessment experts have observed, good assessment practices cannot be distinguished from good learning activities. The process of assessing a complete investigation involves by definition cooperative effort among peers. A component of the peer interaction is giving and receiving criticism of one another's work.

**TABLE 3. PRODUCTS, ABILITIES, AND INFORMATION ASSESSED IN THE PRECURSOR PHASE OF LABORATORY-BASED INVESTIGATIONS**

PRODUCTS	ABILITIES	INFORMATION
Question that will guide the investigation	Obtain and analyze information for the purpose of formulating a question to guide an investigation	About conducting searches for information About how to assess the quality of information sources
Rationale for the question	Formulate questions appropriate to scientific investigation	About scientific inquiry: <ul style="list-style-type: none"> <li>▶ Distinguish among terms related to scientific inquiry (for example, variable, control, experiment, hypothesis)</li> </ul>
Critiques of peers' questions and rationales	Construct a coherent argument in support of a question	<ul style="list-style-type: none"> <li>▶ Characteristics of well-formulated questions appropriate for scientific inquiry</li> </ul>
Reflective report documenting the evolution of the question	Communicate the questions and rationale with peers	<ul style="list-style-type: none"> <li>▶ Practical characteristics</li> </ul>
	Respond reasonably to criticism by peers	About the natural world: <ul style="list-style-type: none"> <li>▶ Scientific facts and principles</li> </ul>
	Judge the scientific quality of a question	About the characteristics of well-structured scientific arguments
	Criticize questions and rationale of peers	About social, technological, civic aspects of the personal context of the question guiding the inquiry
	Keep written records documenting the evolution of the question and the rationale.	

Tables 3, 4, 5, and 6 summarize the products, abilities, and information assessed in each of the four phases of an investigation based in a laboratory. In each phase, the student produces questions, plans, and reports and is responsible for criticizing the products of others.<sup>7</sup>

Tables 3–6 provide a framework for assessing investigations but do not include the performance standards for the abilities and information. After examining the abilities and information called upon in the successful performance of investigations, however, we describe in the next section of this chapter how teachers and curriculum specialists need to work together to develop performance standards for students in grades K-12.

**Precursor Phase.** As shown in Table 3, during the precursor phase students formulate a preliminary question and a rationale for it: a discussion that relates the question to some experience the student has had and an explanation for why the question is appropriate for scientific investigation. Students present to their peers the draft question and the rationale. The peers have the responsibility for commenting on the question and the rationale. On the basis of peer review, students revise questions and rationale.

Work produced by the student during the precursor phase includes the question the student has formulated to guide the investigation. It contains also a rationale for that question: documentation of its origins, a description of its context, an account of the scientific information relevant to it. A reflective record documenting the evolution of the question is called for; so is the student's criticisms of the questions of peers and their rationales.

Among abilities involved in the precursor phase is that of obtaining and analyzing information for the purpose of formulating a question appropriate to scientific investigation. The student must be able to present a coherent argument in support of a question, communicate the question and rationale to peers, respond to criticism, comment on the questions and rationales of peers, and keep written records documenting the evolution of the question.

These abilities, in turn, require the students to have a large store of information and the mental capacity to process it in scientifically acceptable ways. The information necessary for successful engagement in laboratory investigations is considerable. The student must have information about conducting searches, knowledge about the characteristics of good sources of information, and knowledge of scientific inquiry and of characteristics of well-structured scientific arguments. Knowledge of scientific facts and principles must be supplemented by information about the social, technological, civic, and personal context of the decision that the student seeks to make. The student

must have not only a vast store of information but also the cognitive capability to process it in scientifically acceptable ways.

**TABLE 4. PRODUCTS, ABILITIES, AND INFORMATION ASSESSED IN THE PLANNING PHASE OF LABORATORY-BASED INVESTIGATIONS**

PRODUCTS	ABILITIES	INFORMATION
Plan	Collect information from reliable sources to inform the development of a plan	About scientific inquiry
Rationale for the plan	Develop a plan for an investigation (including appropriate scientific apparatus)	About characteristics of investigations
Comments on the plans of peers	Explain how the plan meets the quality standards related to science inquiry	About characteristics of investigations whose methods are well matched to the question under investigation
Reflective report documenting the evolution of the plan	Communicate about the plan with peers in written and oral form, and in sufficient detail that another student could do the investigation	About the characteristics of investigations that meet scientific standards
	Develop a well-structured argument	About scientific equipment and instruments
	Revise the plan on the basis of peer critique	About the appropriate use of equipment and instruments
	Criticize the plans of peers	About safety precautions

Planning Phase. Table 4 shows that during the planning phase students in consultation with their peers develop plans for the laboratory investigation and comment on the plans of their peers. The products produced by students during this phase include the plan, a rationale for it that indicates how it meets the standards of an inquiry related to science, criticisms of their peers' plans, and a reflective report documenting the evolution of the plan.

The abilities assessed in this phase are numerous and complex and much like those assessed in the precursor phase. The major difference is that the abilities require knowledge about the characteristics of well-designed laboratory investigations and scientific apparatus. The application of this knowledge to the plan requires considerable cognitive capability. The ability to develop reasoned arguments, to communicate with peers, to comment on experiments, and to modify work in response to criticisms offered by others are assessed in the planning phase.

Implementation Phase. As shown in Table 5, during the implementation phase students set up the laboratory investigation, collect and analyze data, and develop conclusions from their analysis. As in the previous phases, students consult with their peers to learn how well they have made the logical connections between their data and their conclusions. The products of the implementation phase include the students' data, the data analysis, their conclusions drawn from the data, the argument leading from the data to the conclusions, critiques of the work of peers, and a reflective record of the events and conclusions from the implementation phase. The abilities, skills, and knowledge applied during this phase are different from that applied in the earlier phases in being related primarily to the nature, collection, and analysis of data.

Closure and Extension Phase. The phase sketched in Table 6 has students apply the information they collected in the laboratory investigation to make the practical decision. At this time, information and factors that are related to the context but not to science are brought into the investigation process. Often the decision is not clear cut. In the example of breakfast cereal, economics and taste preference may come into play and even override the scientific information collected by the student. The products of this phase are the students' reports of the investigation and commentaries on other students' reports. The abilities used during this phase are related to the forms of acceptable scientific communication, as well as the reasoning structure of convincing arguments.

As students engage in the investigation, the teacher collects a variety of information. The written products produced by the students, supplemented by observations the teacher makes of students as they interact with their peers and

perform laboratory work, provide evidence about the extent to which students have met the inquiry goal.

**TABLE 5. PRODUCTS, ABILITIES, AND INFORMATION ASSESSED  
IN THE IMPLEMENTATION PHASE OF LABORATORY-BASED  
INVESTIGATIONS**

PRODUCTS	ABILITIES	INFORMATION
Data	Use laboratory equipment and measurement devices	About laboratory techniques
Data analysis		About proper use of laboratory equipment
Conclusions drawn from the data	Make accurate measurements	
	Choose appropriate representation of data	About acceptable forms for data tables and graphs
Argument leading from the data to the conclusions	Interpret data appropriately	About significant figures
	Report data in properly formatted tables and graphs	About measurement units
Comments on the work of peers	Construct an argument developing reasonable conclusions based on data	About measures of central tendency—their calculation and appropriate use
Reflective record of the events and conclusions from implementation phase	Criticize work of peers.	About multiple repetition of empirical investigations

**TABLE 6. PRODUCTS, ABILITIES, AND INFORMATION ASSESSED  
IN THE CLOSURE AND EXTENSION PHASE OF LABORATORY-  
BASED INVESTIGATION**

PRODUCTS	ABILITIES	INFORMATION
Report	Weigh scientific information against	About acceptable forms of reports and scientifically
Criticisms of reports by other students	other kinds of information in making a personal, social, civic or technological decision	acceptable arguments.
	Write a well-reasoned report	
	Develop a complete description of an investigation	
	Develop an argument leading from the original question through the data collection and analysis to the decision	
	Devise a well-structured argument in the application of conclusions from the laboratory investigation to the civic, social, or personal context	
	Consider alternative ways in which the conclusions from the laboratory investigation might be applied in a particular context	

## SETTING PERFORMANCE STANDARDS

Setting performance standards for inquiry is the most challenging task faced by science teachers. It involves consensus between teachers and other district personnel responsible for the science curriculum. The framework for investigation in Tables 3–6 guides the process of setting performance standards for investigations, whether the laboratory or other sources of information are at issue. The performance standards must be set for students at each grade level. The preliminary step in a process for setting performance standards that we have found successful brings together groups of teachers who develop descriptions of performance that would convince them that students at their grade level are progressing toward meeting district standards. The teachers then design a task that will elicit the performance from their students. Teachers collect samples of resulting student work and match these against their expectations. In collaboration with their colleagues, they reach final consensus on reasonable performance standards for students at the grade level they teach.

For instance, teachers at the sixth grade would need to set performance standards for investigations by students at that level and the quality of students' explanations for why their plans are scientifically sound. Working in grade level committees, teachers would develop descriptions of the characteristics they expect for their students' investigations and explanations and then decide how they will assess their students' abilities to meet their expectations. After students have had ample opportunity to learn to design investigations, teachers would have students design plans. All sixth-grade teachers in a school or district would bring their students' plans and explanations to a meeting, where they discuss how well the work matches the expectations. Through a process involving the refining of the expectations in light of student work, performance standards for plans and explanations for the sixth grade would be developed. To achieve articulation for this ability and knowledge across grade levels, teachers from all grades need to work together to compare expectations as a means of identifying consistencies, redundancies, and inconsistencies in characteristics they have identified.

The process we have described is time consuming, but we believe that teachers who engage in it will come to a deeper understanding of what it means to inquire and the challenges of teaching students to inquire.

**CLASSROOM-BASED ASSESSMENT**

To plan and evaluate their classroom practices, teachers use the information they collect in their classrooms as summarized in Table 7. Teachers gather it week by week, day to day, minute by minute. They use this information to make judgements about the effectiveness of teaching methods and decisions about which methods will be used. The most recognized use of the information collected in class is to monitor and report student progress, especially grading students and reporting their achievement to parents.

**TABLE 7. CLASSROOM-BASED ASSESSMENT**

PURPOSE	DECISIONS	ASSESSMENT STRATEGY	RESPONSIBLE INDIVIDUALS
Inform Classroom Practice			Committees of teachers and students responsible for the course  Individual teachers
Daily planning	To make on-the-spot modifications in teaching plans To provide individual students special instruction	Monitoring Student Learning Teacher gathers information on a minute-to-minute basis about students' understanding and abilities <ul style="list-style-type: none"> <li>▶ Questions students during lessons</li> <li>▶ Cursory reviews of student work</li> <li>▶ Short-term observations of student performance</li> </ul> Monitor Teaching strategies	
Weekly and longer-term planning	To make extensive modifications in teaching plans	Teacher gathers information on daily or weekly basis about students' understanding and abilities <ul style="list-style-type: none"> <li>▶ Gives short quizzes</li> <li>▶ Reviews student work</li> <li>▶ Makes long-term systematic observations of student performance</li> </ul>	

Testing the effectiveness of teaching methods	To select effective teaching methods	Measures of student achievement Measures of opportunity to learn
Reporting Student Progress	What grade to assign the student	Measures of student achievement
Grading Reports to parents	Whether the student will pass or fail the grade	

Some decisions made by teachers using the information they collect have profound and long-term effects on the lives of students and teachers. Such decisions must be based on information in which teachers have great confidence. For instance, decisions to promote have a more profound effect on the lives of students than decisions to modify a student's daily assignment. Decisions about assignment changes can be easily rectified in light of new and better information without serious consequences. Such is not the case for promotion decisions.

The quality of educational information is determined in large measure by its accuracy and congruence with the decisions that will be made using it. If inquiry is a central objective of a course, grades in the course should be based on information about students' ability to inquire. Grades based on information gathered using multiple-choice tests are suspect because the information is not congruent with the course objectives.

The role of teachers in the assessment of inquiry cannot be confined to the grade level they teach or to assessment made in the classroom. They must also be involved in the design and implementation of the district's overall plan for inquiry in the science curriculum. Assessment plays an integral part in the planning process at the district level.

## ASSESSMENT IN PROGRAM AND COURSE PLANNING

This chapter is predicated on the assumption that the development of effective assessment strategies for inquiry is a collaborative process. At a minimum, science curriculum specialists and science teachers must work together to develop assessment plans that include methods of data collection and analysis. Expectations for proficiency in inquiry that all students should attain, as

well as strategies for the assessment of inquiry, are too poorly defined and their development too challenging for individual teachers working in isolation to accomplish.

Developing such strategies is an integral part of K-12 planning for science programs and courses. Program and course planning is facilitated by the development of assessment strategies at the beginning of the planning process rather than treating assessment like the little red caboose that always came last. Definition of a science program's goals and objectives in measurable terms—an essential step in the assessment development process—facilitates communication among program planners about the goals and objectives and enhances the reaching of consensus. Such consensus is the essential precursor to further steps in the program planning process: development of scope and sequence, selection of pedagogical practices, and textbook choice. Definition of the goals of a program or course and objectives in measurable form not only facilitates further program planning but also is the beginning to the design of the assessment plan for the program.

If a program goal is to teach its students to inquire, members of the program planning team are challenged to reach consensus on an essential question: what evidence will they accept that their students have met the goal? Answering this question requires that individual program planners agree on the nature of the evidence that will be collected and how that evidence will be interpreted. The purpose of the assessment plan is to decide whether the district's students have met the program goals. The assessment planning process requires consensus among the responsible individuals about how the data that will be used as evidence will be collected.

The assessment plan should question not only the achievement of students but the extent to which the school science program provides the students with the opportunity to develop the knowledge and abilities required for proficient inquiry.<sup>8</sup> The opportunity to learn question is much the same in form as that of student attainment: what evidence will we accept that we have provided our students adequate opportunity to develop the abilities of science inquiry? Designing the strategies that will be used to collect the evidence is integral to the process of curriculum planning. Reaching consensus on the strategies for assessment of opportunity to learn ensures that the responsible individuals are in agreement on resources that need to be in place if the student attainment goal is to be met.

Table 8 summarizes purposes, decisions, assessment information, and responsible individuals for program and course planning and evaluation.

**TABLE 8. ASSESSMENT IN K-12 PROGRAM AND COURSE PLANNING AND EVALUATION**

PURPOSE	DECISIONS	ASSESSMENT INFORMATION	RESPONSIBLE INDIVIDUALS
<b>K-12 Program and Course Planning</b>			
Reach consensus on program or course objectives	Which of the possible program or course objectives will the K-12 program select?	Definition of program or course objectives in measurable form: design of prompts and performance expectations	Program Planning: Committee consisting of teachers, administrators, curriculum specialists, parents, community representatives, and students
Reach consensus on the conditions under which students can meet the objectives	Which teaching materials and strategies to select?	Definition of the conditions in measurable form	Course Planning: Committee of teachers (and students) responsible for the course
Identify resources that will be necessary to meet the conditions	Does the district have the resources to meet the conditions?		
<b>K-12 Program and Course Evaluation</b>			
Determination of how well the existing program or course is meeting the district's expectations	Whether the program or course will be continued or requires modification	Measure students' attainment of the program or course objectives: use the prompts and performance expectations developed in the program or course planning phase to measure and evaluate the effectiveness of the program or course	Program Planning: Committee consisting of teachers, administrators, curriculum specialists, parents, community representatives, and students
		Measure the extent to which the conditions necessary for students to meet the program or course objectives were met	Course Planning: Committee of teachers (and students) responsible for the course

### Program and Course Planning

As indicated in Table 8, consensus on program goals and course objectives and the classroom conditions is essential to planning programs and courses. Assessment plays a central role in the process of reaching consensus on these matters. Definition of the objectives in measurable form, that is, specification of the prompts and performance expectations for the objectives, is an effective way of improving communication among the individuals responsible for making the decisions.<sup>9</sup> The planning process also requires consideration of the resources necessary to provide students with adequate opportunity to achieve the objectives. Assessment, in this case, defining in measurable terms the resources required to provide that opportunity, is an effective mechanism for clarifying the necessary resources. Once the requisite resources have been identified, the responsible individuals have the information necessary to decide whether the district can provide the resources.

### Program and Course Evaluation

Program and course evaluation is a generally recognized purpose for assessment, and is therefore included in Table 8. But far too often evaluation planning comes after the program or course is developed and implemented. When objectives and resource requirements are defined in measurable terms in the program and course planning processes, designers know how the program or course will be evaluated. Such information guides the development process in ways that increase the likelihood of a favorable evaluation.

### CONCLUSION

This chapter has shown assessment of inquiry to be a complex process, intensive in the need of resources. The principles of good assessment are more easily defined than implemented. The *Benchmarks* and *Standards* can provide guidance in defining goals. They do not, however, give details on how a school district, classroom teacher or student of inquiry will actually meet those goals. If inquiry is a district goal, the district must be ready to devote considerable resources to providing students with adequate opportunity to meet that goal, to supplying teachers with adequate occasion to develop strategies to teach inquiry abilities, and to assessing attainment of the goal of inquiry.

## ENDNOTES

1. Both documents call for students to become knowledgeable about inquiry as it is practiced by scientists and for engaging students in inquiry to develop their understanding of the natural world. The *Benchmarks for Science Literacy* (AAAS, 1993) place greater emphasis on students knowing about inquiry than on their acquiring the ability to conduct inquiries. The *National Science Education Standards* (NRC, 1996) place greater emphasis than the *Benchmarks* on students' acquisition of the ability to conduct scientific inquiries (Kouba & Champagne, 1998).
2. For psychologists the terms "learning" and "development" have different meanings. The terms as they are used in this chapter emphasize the complexity of the knowledge, abilities, and inclinations possessed by the competent adult inquirer and acknowledge that the requisite knowledge and abilities are acquired over a long period. Consequently, as we speak of inquiry in the context of the K-12 program, we used the term "development." When referring to the individual components of information and the many individual abilities that may be acquired in relative short periods, we speak of "learning."
3. In education the terms "program" and curriculum" are used interchangeably. We have chosen to use "program," which we define to include goals, content, preferred teaching methods, and assessment strategies; as "objectives" contain detailed and explicit statements about what students are expected to learn and the order in which students are introduced to the content. Typically a curriculum contains, in addition to the elements of a program, statements about its philosophical, psychological, social, political, and disciplinary foundations.
4. For a detailed discussion of how assessment can be used by curriculum developers to match curricula with standards, see Champagne (1996).
5. The *Benchmarks* (AAAS, 1993) and *Standards* (NRC, 1996) call for schools to produce students who are literate in science as well as for the development of the abilities of science inquiry. But it is not clear whether the ability to inquire is equivalent to being literate in science (Kouba & Champagne, 1998).
6. Our placement of experimentation in high school is consistent with its placement in the *Standards* (NRC, 1996).

7. As outcomes for science education, the *Benchmarks* (AAAS, 1993) and *Standards* (NRC, 1996) include cognitive processes such as scientific reasoning, logical reasoning, critical thinking, analysis, and evaluation. Cognitive processes cannot be observed directly. Teachers must base their inferences on samples of a student's work or observation of performance. Our analyses draw on particular models of cognition, such as those proposed by theorists of information processing (Simon, 1962; Bloom et al., 1956; Inhelder & Piaget, 1958; or Gagne, 1970). Assessing cognitive process is challenging because it is so inferential and dependent on theory. Consequently we have not included in our discussion of the assessment of laboratory inquiry descriptions of the cognitive process that are the foundations of the abilities of the proficient inquirer.
8. The *Benchmarks* (AAAS, 1993) and *Standards* (NRC, 1996) advocate inquiry as a pedagogical method to develop understanding of the natural world and the character of scientific inquiry. Because understanding both contributes to the ability to inquire, assessing the degree to which students are exposed to inquiry speaks not only to the abilities of inquiry but also to the knowledge required to be a proficient inquirer.
9. In defining the concept of inquiry, we follow the suggestion that Alfred North Whitehead made in 1925 that scientific concepts must be defined in terms of their measurement.

## REFERENCES

- American Association for the Advancement of Science. 1993. *Benchmarks for science literacy*. New York: Oxford University Press.
- Bloom, B. S., M. B. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. 1956. *Taxonomy of educational objectives. Handbook I: Cognitive domain*. New York: David McKay.
- Champagne, A. B. 1996. Assessment and science curriculum design. In *National standards and the science curriculum: Challenges, opportunities, and recommendations*, edited by R.W. Bybee, 75-82. Colorado Springs, CO: BSCS.

Kouba, V.L., and A.B. Champagne. 1998. *Literacy components in natural science and mathematics standards documents: Communication and reasoning*. Albany, NY: National Research Center on English Learning & Achievement.

Gagne, R.M. 1970. *The conditions of learning*. New York: Holt.

Inhelder, B., and J. Piaget. 1958. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.

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

Simon, H. A. 1962. *An information processing theory of intellectual development*. Monographs of the Society for Research in Child Development 27 (2, Serial No. 82).

Whitehead, A.N. 1925. *Science and the modern world*. New York: New American Library, The MacMillan Company.

# Implications for Teaching and Learning Inquiry: A Summary

Jim Minstrell

---

## INTRODUCTION

When Emily van Zee and I were invited to edit this book on inquiry, I accepted because I thought the process and product would help me better understand the nature of scientific inquiry and how it might be taught and learned. We got to choose the authors. Although there were others whose work and ideas we would like to have included, the set here makes up a very good team in having both diverse and overlapping ideas. There are some minor differences of opinion, but on the whole, the chapters help to create a coherent vision of many attributes of inquiry.

The first section of this chapter presents my brief summary of some aspects of inquiry I see exhibited in the chapters in this book. The second presents through a vignette a vision of what inquiry might look like in my classroom. In my short space, there is no way I can do justice to all the efforts and ideas expressed by the other authors. I strongly suggest that you convene a group of educators to have discussions around each of the chapters and come up with your own summary, and then work together on implementing your vision in the learning environments you create.

## BRIEF SUMMARY OF LESSONS ON INQUIRY

### Inquiry Seems to Mean Different Things to Different People

At the present, we may agree more easily on what inquiry isn't than on what it is. We seem to know when what we are seeing is not inquiry. Perhaps we can know inquiry by knowing what intentions it serves. Is the central purpose of your inquiry to prepare students to develop new knowledge or to contribute to a satisfying and productive life? What goal we choose determines how we instruct and how we assess the achievement of those learning targets.

What does it mean to do inquiry? One definition is that it involves fostering inquisitiveness; inquiry is a habit of mind. Another holds it to be a teaching strategy for motivating learning. But what does that mean? Some say inquiry means "hands-on." Others say that is too simple: that hands-on is not necessarily minds-on, and inquiry requires mental reflection on experiences. Some educators add that inquiry includes manipulating materials to become acquainted with phenomena and to stimulate questions, as well as using the materials to answer the questions. Scientific inquiry is a complex process. There is no single, magic bullet. Some instructors, in the interest of a focus on inquiry, address only one aspect of the complex process. That contributes to the impression that inquiry means different things to different people. So we need to identify the various aspects of the process and see them as a whole.

We need to encourage and support personal curiosity when it occurs spontaneously and stimulate it when doesn't occur naturally. Then we need to challenge students to do deep thinking and find answers for themselves. When learners are involved in the struggle to understand something for themselves, they are more likely to feel the sense of pleasure when the understanding does come. I'm reminded here of the intellectual pleasure and sense of personal accomplishment I felt when I brought myself to deeper understanding with only the assistance of questions from my mentor to help guide my personal investigation. People like to confront the unknown. They are motivated to understand what they presently do not. In general, we have not taken the time to understand students' preconceptions; we give their ideas little credit or status. But for students to engage in inquiry, the questions and ideas being tested best come from them, or we run the risk of continuing to lose their intellectual interest.

Scientific inquiry is not new. For some time science educators have been trying to define and emulate the examples of inquiry conducted by early scientists. My favorite is Galileo's "Dialogues Concerning Two Chief World Systems," which gives a glimpse of scientific inquiry in the interaction among three

discussants. Clearly one lesson is that scientific argument is about making sense of observations, not standing on the opinions of authority. When we have finished the inquiry we should know something we did not know before we started. We should be able to make more accurate predictions about events in nature. Even when our investigation fails to find the answer, at least the inquiry should have yielded a greater understanding of factors that are not involved in the solution.

What are the assumptions behind learning and teaching inquiry? Here is my summary of assumptions that follow from the other chapters of this book.

### Almost Anyone Can Learn by Inquiry

Practicing scientists and young children, nearly everyone can learn from inquiry. A common interpretation of the work of Piaget is that young children cannot do inquiry, for they are not capable of formal, hypothetical, and deductive reasoning. But several of our authors are quick to point out that Piaget's work describes uninstructed children's capacity for engaging in the reasoning involved in scientific inquiry. When children have an opportunity to learn in a scaffolded but challenging reasoning environment offering evidence and built from personal observation and experiences, they show that they can reason at much higher levels than even the national standards documents suggest. Several chapters have described special populations, such as inner city, economically disadvantaged children or children with various physical or learning disabilities (with the possible exception of mentally retarded) learning by inquiry, as well as they learn by more traditional methods or better. But that requires specific teaching practices and supportive curriculum.

### Inquiry Should Be Incorporated into Planned Learning Goals and Teaching Methods

We need to improve learners' knowledge about inquiry and their ability to do it. Past approaches to learning will not be good enough for the future, for this new century. Scientific information and procedures can be helpful in making day-to-day personal, social, and economic decisions. In general, people will need to be able to ask questions and seek and evaluate information that may help them answer those questions. Of course developing scientists will need to be able to conduct inquiry. But the general population can also profit from scientific inquiry that will satisfy intellectual curiosity or enhance the quality of life as well as solve practical problems. Some say inquiry when properly done can motivate life-long learning.

Schools will likely need to make some changes to meet this future. Many of today's students are disengaged with school activities. As several of the authors in this volume suggest, colleges and universities need to reform their practices as well. Teachers and others can benefit from courses that are challenging. Even recent science Ph.D. graduates drafted to become pre-college science teachers will need to learn about the nature of learning and understand the relation between learners' present knowledge and the sort of experiences that can help foster growth to a deeper, more principled scientific understanding.

National and many state standards documents recommend these kinds of learning connected with inquiry: learning about inquiry, learning to do inquiry, and learning by inquiry. While learning about inquiry can be through studying the work of others, the last two involve active participation. But even if learning subject content is by active involvement in inquiry processes, students may not learn to use the processes of inquiry unless teachers explicitly target and address the learning of specific aspects of inquiry. More than one of the authors point out that teaching both subject matter and inquiry processes at the same time most frequently brings students to learn one or the other but not both. Each goal of learning needs specifically to be identified, taught, and assessed.

### The Inquiry Process Is Motivated and Sustained Through Dialogic Questioning and Reflection

Although the authors differ on whose question initiates the inquiry, all agree that the learner must own the questions around which the inquiry takes place: that the questions are understandable and intellectually interesting to the learner. Some of the teacher authors give examples of how they generate interest in the questions for the learners. At the elementary level, sometimes the interest can come from stories in which the children have already shown curiosity. Other times we have to work a bit harder and relate our questions to aspects of their daily experience.

As a teacher I worked hard to formulate questions that both addressed content important to physical science and yet were interesting to the general high school population. For example, to get kids interested in questions of measurement, I designed a scenario involving multiple measurements of blood alcohol for a particular driver suspected of driving under the influence. "What should be reported as the blood alcohol count for the driver?" "Why would there be different readings from the same sample of blood?" "So should the driver be presumed to be under the influence or not? How did you decide?"

These questions, while I initiated them, got nearly all my high school juniors and seniors mentally actively involved in an inquiry about the nature of measurement and related issues. Whenever during the rest of the year we again encountered issues of measurement, students referred to what they had learned from this inquiry.

In a community of inquiry, whether in the classroom or in the informal environment of a science museum, the learning of each member is affected by other learners and guides as well as by the physical environment. Carrying on a dialogue with one's self and with others stimulates deeper reflection. As we share our observations, questions, and ideas with others, we gain clarity of meaning for ourselves. Many of the authors focus on creating cooperative or collegial arrangements for learning with and from fellow learners. But these environments are not easy to manage. Learners need to be guided to probe their own thinking and the thinking of others.

The verbal interaction between teacher and students and among students needs to embody respect for one another at the same time it is critical of the ideas being expressed. When learners share their questions and their ideas for making sense, listeners foster growth by helping the speaker to express and represent his or her ideas clearly. Appropriate scaffolding includes respectfully helping the speaker identify possible reasoning errors or a need for more experimentation to clear up ambiguities in the data and to support an argument. "Metacognitive reflection" on the part of the individual learner is enhanced by support from the learning community.

### **Inquiry Includes Reasoning From Experience and From Ideas That Describe and Summarize Experience**

Virtually every chapter talks about mental activity on the part of the learner involved in scientific inquiry. Reasoning doesn't happen in a vacuum. Experience with materials, phenomena, and ideas provide the grist for reasoning in science. In the process of generating new ideas, they are induced from experiences with phenomena. Predictions are reasoned from phenomena or from previously generalized experiences. Conclusions are generalized from results. When learning about scientific inquiry, students need opportunities (and scaffolding) to separate their observations from their inferences. Several of the authors point out the common difficulty of separating conclusions (which require personal reflection) from results (which are organized observations of the experience). To stimulate reasoning from experience, critical questions to

ask learners include “What happened, what did you see or hear?” (What did you observe?) and “Why do you think that happened, what sense do you make from these observations?” (What do you infer? What is nature trying to tell us?)

Although many of the chapters suggest hands-on experiences, most authors are quick to point out that hands-on should not be equated with inquiry. Reasoning about experiences that have become personally viable to the learner is what matters. The experiences might have been initiated by hands-on or by a shared, reliable (and probably repeatable) experience by others. First, learners need to “know” the phenomena of the experience. Doing it or seeing it for themselves can help. Then, they can be stimulated to reason to make sense of the experience.

In science it is important always to couple reasoning with experiential evidence. Observations and related inferences of meaning are primary intellectual activities. Opinions are not a necessary part of science. Personal opinions may be formed as a result of scientific inquiry, but opinions can also arise out of personal whim, what a learner wants to happen or dreads might happen. A scientific inquiry goes on to ask for the experiences that suggest that that phenomenon might happen and the reasoning that leads from those experiences to a conclusion. Or opinions may be formed from having done the science. The scientific inquiry may have provided the reasoning behind the formation of an opinion.

### The Materials Used in an Inquiry Environment Need to Support the Inquiry Learning Expected

The worksheets, the laboratory or demonstration equipment, the software, the text resources, and assessment activities all should support the inquiry goals of the program. For example, the printed materials can invite and guide thinking in what to think about, but should not short-cut thinking by directing students what to think. Materials can be made available to learners after questions have been initiated by learners, or materials and minimal instructions for what to do with the materials can be provided to stimulate more questions and thinking. Software can assist students in recording their results and can scaffold students in the development of conclusions from results without directing students in what they should think. Text and technological resources of information can be useful to learners, when they have the questions for which they want answers. These resources can assist learners in making more efficient the search for specific information. Assessment tasks, whether on paper or served through technology, can help learners and teachers identify a needed relevant experience or a weakness in the learner’s reasoning from prior experiences. Assessment of inquiry

skills, aligned with the specific inquiry goals of the educational program, must be as explicit as assessment of understanding of ideas.

Meanwhile, the tone of the environment by teachers, software, or demonstration medium needs to be supportive of thinking itself rather than set by extrinsic rewards for getting the right or punishment for arriving at the wrong answer. If learners obtain results that are not consistent with what others are getting, that needs to be identified. Then, help learners identify why their results may be different and encourage replication. If conclusions are not deemed as correct, help the learners identify inconsistencies between their thinking and prior experiences. Frequently teachers may find that it is the curriculum materials that are lacking rather than the students' thinking.

The point is to create an atmosphere of identifying and supporting learning needs rather than dictating what learners are to do and think. Setting a proper set of experiences before the learners will enable them to make their own good sense of the experience.

### Learning to Do Inquiry and to Teach Inquiry Are Best Done Through Experience in Learning by Inquiry

Learning itself is an active process, and inquiry is best learned through actively participating in the process. Children should have opportunities to learn to do inquiry by participating in the inquiry process in class, or in the informal setting, such as the museum or group or family experiences in natural environments. The same is true of teachers, whether in elementary or in secondary schools, and the principle applies to young scientists switching careers to teaching. Since teachers typically teach as they have been taught, learning by inquiry will increase the probability that they will teach by inquiry.

Teachers also need opportunities to engage in their own professional inquiry into the teaching and learning of their students: to talk with others about what sorts of conceptions and procedures students are likely to develop, what kinds of instructional activities will dissuade students from inquisitive ideas and skills, and what sorts of assessment activities will allow the teacher and students to monitor development. Teachers need opportunities to conduct teaching experiments to understand better their students' learning.

## PUTTING IT ALL TOGETHER

Many of the chapters give examples of the authors' views of inquiry in the classroom. I can't resist the opportunity to share an example from my experience. While I can't possibly exhibit all the ideas presented in this book, I would like to address some of the critical assumptions and issues expressed in the prior section, particularly the concern about trying to do many things without cheating on any of them. How is it that a teacher might concentrate on teaching by inquiry both specific content and specific aspects of inquiry? How might a teacher create and maintain an environment for inquiry while centering on development of specific concepts and skills? In this vignette, I will give a flavor for how I might teach about and by inquiry and have students develop their conceptual understanding as well.

### The Context for This Vignette

We are about to begin a unit on forces and explanation of motion, after which I expect that the students will explain various motions, using ideas more consistent with Newton's Laws. But I want them also to gain a better understanding of scientific inquiry and skills for doing it.

### Raising Questions

At the beginning of a unit I find it helpful to pose questions that survey students' present understanding of ideas or situations that will be the source of our class inquiry for the next several days or weeks. As the students construct their answers individually, they start to generate their own questions about the issues of the unit, and they become motivated to answer questions they and others have initiated.

From research on students' understanding, I know that one of the alternative explanations for motion that makes sense to them is something like: "to explain motion, one needs a net force, acting on the object. The size (magnitude) of the net force should be proportional to the speed of the object" from their point of view. I call such constructions of ideas "facets of students' thinking" (Minstrell, 1992). Opportunities for students to express these facets of their thinking need to be built into my initial elicitation questions.

A first question might involve pushing a book across the table top and asking students to use arrows to represent each of the forces acting on the book while it is moving with a constant velocity, and it includes asking students to clarify what

is exerting each force. A follow-up question asks students for diagrams that represent the forces on the book as my hand pushes and the book accelerates across the table. For each case, I ask for five diagrams at equal time intervals. Students want to talk about getting the object going and stopping it at the end. The three images I require for the middle of each motion story gets them to think instead about the constant velocity or the constant acceleration (Minstrell, 1984).

### Determining the Hypotheses

Following the pre-instruction questions, I conduct a full-class discussion. As students share their answers and ideas, I write the ideas on the front board for all to see. In this case the two events we are trying to explain consist of constant velocity of an object and its constant acceleration. Later we will worry about more exotic or unusual motions. The answers and ideas evolve into initial ideas (hypotheses) to be tested through classroom laboratory experiences.

In this case for the constant velocity in a straight line, typically the class suggests these two ideas: A constant extra (unbalanced) force is necessary to keep an object moving with a constant speed (stated in “if.., then...” form, it would be “An object moves with a constant velocity if, and only if, there is a constant net force acting on the object”); and Once the object is moving, no (zero) extra force is required to keep it moving with a constant speed. For the constant acceleration cases the two typical hypotheses run something like these: An extra force that increases at a constant rate is necessary to keep the object accelerating at a constant rate; and A constant extra force will be required for the object to accelerate at a constant rate. In my classes, typically in excess of ninety percent of the students believe the first hypothesis for explaining constant velocity and the first hypothesis for explaining constant acceleration. About five percent suggest the second hypothesis for explaining constant velocity and the second hypothesis for explaining constant acceleration. Occasionally another hypothesis is listed.

“Are the ideas we’ve written on the board true?” I ask the students. While some initially say yes, others are quick to point out that they can’t all be true, because they conflict with each other. After some discussion among students, they suggest that each statement is an idea that somebody in the class thought was true. After students have articulated a recognition that the suggested explanatory ideas are tentative, I observe that in science these testable ideas we think might be true are called “hypotheses.” Thus, I write “Class Hypotheses” above the set of ideas we are about to test. But before I used the more technical term “hypothesis,” I had the students articulate the notion of a tentative (and testable) idea.

I usually run the survey and the preceding discussion in about twenty minutes at the end of one class period. I may ask students to consider (and record in their journals) various situations in their lives that seem to support one or another of these hypotheses. The next day I have the initial class hypotheses written on the board and the discussion continues.

### Identifying Relevant Variables

“So if we are going to do an experiment to determine what idea best fits careful experimentation, what variables are we going to need to measure?” With some discussion the students decide that the relevant variables are speed and force. Some students are considering force as simply the pulling or pushing force by whatever is making the object move forward. Others are concerned that we need to consider any retarding forces like friction as well. Finally, it is decided that the extra force (net force) is the important variable, and that will involve a resolution of any backward forces as well as the forward force on the object.

### Simplifying the Experiment

In subsequent discussion, some students voice concern that the backward force of friction might be difficult to measure. Preferring to avoid the topic of friction at this point in our class inquiry, I guide the students in setting aside this concern. Several students readily suggest we design the experiment so we have no (or negligible) friction. They choose to do the experiment using a low friction cart as the object that will move rather than a book sliding across the tabletop.

### Measuring the Relevant Variables

“What measurements will we need to make?” The students suggest that if the cart wheels are exerting very low friction, we only need to measure the force of the string pulling the cart. That will be the extra force causing the motion. “Any other measurements we will need?” Some students suggest we will need to measure the motion. “How might we do that? What do you mean by motion here?” Students suggest we will need to measure the speed of the cart. “What speed? The average speed or the instantaneous speed?” Some students suggest we just need to get the average speed because we are looking for how big the speed is in relation to how big the pulling force is. “Is that right, do we know the speed will be constant? For sure?” Other students are quick to point out that that

is the purpose of our experiment, to see what sort of motion is related to the extra force. They suggest we do what we had done during earlier studies of kinematics, the description of motion: measure several positions and the corresponding clock-readings (times), during the motion of the cart across the table. From those data, they suggest, we will be able to construct a graph of speed against time and describe the motion. Measurement and kinematics were among our earlier units, so they are familiar with issues of best value and uncertainty in measurement of quantities including position and time. In the preceding sub-unit on explaining the condition of the object at rest, we have used spring scales for measuring forces (Minstrell, 1982).

In this case I have used what Emily van Zee and I (1997) have called a “reflective toss,” taking one student’s answer and reflecting it back to that student or to the class as a whole. Rather than having students key on the teacher for knowing what to do, I want the students to be responsible for building their own understanding. As their teacher, I am there to juxtapose related experiences and to use questions to guide their thinking about the experiences and ideas. They need to decide what to do and why and what the results will mean.

### The Experiment

In the back of the classroom I have anticipated and already set out the sort of apparatus the students will need. They have carts with spring scales taped to them. A string is attached to the spring scale and leads down the table around a pulley, to the top of a ringstand, and around another pulley and is attached to a 100 gm mass that hangs out over the end of the lab table. This is a variation of an experiment from *Project Physics* (Rutherford, Holton, & Watson, 1972). An even more primitive version was in *PSSC* (Haber-Shaim et al., 1959/1971). Newer versions are available using micro-computer based force probes and motion sensors, but I prefer that the initial experiment be done with the spring scale, meter stick, and stopwatch, the measurement tools that are more transparent than the computerized probes and are more accessible to students out of school. I want to encourage students to put together experiments at home. Because the movement across the table only takes a couple of seconds, I have had to use a tickertape timer to get a representation of the positions across uniform “dots” of time. Students have used this same apparatus in earlier units.

### Obtaining the Results

With the suggested apparatus, when students let go of the 100 gm mass hanging from the pulley, the cart accelerates across the table leaving a set of dots at regular time intervals on the tickertape. While the cart is moving across the table, students read the spring scale. When they graph the motion it turns out to be a straight diagonal line for speed vs. time. That represents a constant acceleration. While the spring scale reading drops slightly from its initial value, it then stays constant (except for vibration) all the way until the cart is caught. This suggests the constant extra pulling force resulted in a constant acceleration. The students then repeat the experiment with a 200 gm mass and then a 300 gm mass hanging over the end of the table. For these the acceleration is roughly proportionately greater which is represented in a steepening of the slopes of the speed-time graphs, and the spring scale readings are also proportionately greater.

### Inclination to Deny the Results

Although some students just did the experiment and got their results without thinking, others asked whether they could have another spring scale. They suggested that their results had not come out right. They observed that while they were obviously getting accelerated motion—dots were getting progressively farther apart—the spring scale seemed to stick on one number. This, of course, was what I had expected to happen, but I gave them a different spring scale without suggesting that their results were OK. Others came and said they could never do science and get the right answers. Wanting to boost their confidence in their capabilities, but also wanting them to discover that their results were all right, I proposed that they check their results against the results obtained by other groups. An occasional group would excitedly come to share with me their recognition that the dominant class hypothesis was not right. Several recognized their original intuition had not been supported. Students finished their analysis of the data for the three different masses and got ready for a large group discussion of conclusions from these experiments.

## Post-lab Discussion and Conclusions

We assembled for a whole class discussion. Still on the board was written:

Class hypotheses

To explain movement with a constant speed in a straight line

1. A constant extra (unbalanced) force is necessary to keep an object moving with a constant speed.
2. Once the object is moving, no (zero) extra force is required to keep it moving with a constant speed.

To explain movement with a constant acceleration

1. An extra force that increases at a constant rate is necessary to keep the object accelerating at a constant rate.
2. A constant extra force will be required to keep an object accelerating at a constant rate.

Which hypothesis for each motion seems to work?

Here is a paraphrase of typical past discussions:

**Teacher:** So, what do your results tell you about the relation between force and motion? What can you conclude? (I am trying to get the students to move from results to conclusions, a difficult differentiation for many students, as noted by a few authors in this volume.)

**Student:** We got constant acceleration.

**Teacher:** For how many of you did the acceleration turn out to be constant? (Most raise a hand or otherwise acknowledge that result.)

And how did you know the acceleration was constant? (Several students observe that the graphs of speed against time had constant slopes or that the spaces on the ticker tape were getting larger in a regular way. My question provided an opportunity for students to visit, and practice, analyses they had done in a previous unit.)

**Teacher:** OK, but was that the purpose of the experiment—to see whether the motion was a constant acceleration?

**Student:** We were also trying to see what the force would be like.

**Student:** We were trying to see whether the force changed or stayed constant.

**Student:** The force changed. It went down.

**Student:** Well, not really. Actually, it only went down right at the start. Then it stayed constant.

**Teacher:** Who observed that the scale reading dropped off? (Nearly everyone agrees.)

**Teacher:** Did the force scale reading gradually drop off, all the way across the table, or did it just drop down a bit at the beginning and then stay more or less constant the rest of the way before the cart was caught? (Students confer with one another and, after some argument among themselves, agree that it dropped right at the beginning but then stayed constant. One of the groups that wasn't sure goes back to the tables with the apparatus set up and quickly retries a run and announces that it dropped at first and then stayed constant. Another of my goals for interaction in my classroom is that students question one another and argue on the basis of evidence even if it means they need to repeat the experiment to get clear on the results.)

Initially I was trying to get the students to focus spontaneously on formulating a conclusion about the relation between force and motion. They concentrated instead on citing specific observations, so I changed my immediate focus to get them to clarify their observations. But now I do want students to move from results to conclusions, in this case a conclusion about the relation between force and motion.

**Teacher:** Suppose we look back at these initial hypotheses (gesturing to the hypotheses on the board) that represented our class ideas about the possible relations between force and motion. We called these our "class hypotheses." Did these experiments relate more directly to explaining the constant velocity or the constant acceleration?

**Student:** The cart didn't move with a constant speed.

**Student:** The cart was accelerating.

**Student:** The constant acceleration, 'cuz the cart accelerated at a constant rate.

**Student:** The bottom two hypotheses about constant acceleration.

**Student:** But both of those can't be true, because one says an increas-

ing extra force and the other says a constant extra force, and it can't be constant and increasing at the same time.

**Student:** Number 1 can't be true because it says an increasing extra force, but ours decreased and then stayed constant.

**Student:** So number 2 must be the one that is true. We got a constant extra force reading and we got constant acceleration.

**Student:** But I thought we decided that the force should be increasing constantly to get a constant acceleration?

**Student:** Well, that was what we thought to begin with, but that is not what happened in the experiment. So we must have been wrong with our first ideas.

**Student:** Our first hypothesis was wrong. The experiment comes out more like hypothesis number 2.

**Teacher:** OK, does that make sense? When you did your experiments, your results involved getting constant acceleration of some number and you got a constant spring scale reading, while it was moving across the table. Each of you got slightly different numbers for the force and for the acceleration, but in every case the results involved a constant acceleration and a constant extra force. We can now write a general conclusion: "To get an object to accelerate at a constant rate requires a constant extra force." From our experiments, this seems like a reasonable conclusion. Our only other hypothesis got eliminated. Right? Does that make sense? (Most students chime in with the affirmative.) Notice that is different from what most of you thought to begin with. That is OK. Experiments often can tell us that our present thinking is NOT correct. That is valuable to know.

**Student:** I see that. It makes sense, but then, how do we explain the constant velocity? Both motions can't be explained the same way.

**Teacher:** Good next question. But first, are there any other questions about the constant acceleration cases? (After asking a question that requires deep thinking or for which students will need time to formulate their ideas or concerns, it is advisable to wait for something over three seconds [Rowe, 1974].)

**Teacher:** OK, so now let's consider an object moving along at a constant speed in a straight line. How might we explain that? Do these past experiments tell us anything about that? (More wait-time.)

**Student:** (Looking at the board) Well it can't be a constant extra force, because we just said that explains constant acceleration.

**Student:** But didn't we agree yesterday, even before we did the experiment, that constant force makes for constant speed?

**Student:** Yeah, that's what we all, well most of us, thought before we did the experiment. But the experiments proved us wrong.

**Student:** It can't be constant extra force for constant speed in a line, because we just got a constant force and a constant acceleration. Constant force can't do both.

**Student:** Well, it wouldn't make sense to say no force is needed to keep it going with a constant speed. (Actually, eventually I am hoping that the students will see that this is the only valid conclusion given the results.)

**Student:** What else could it be?

**Student:** We eliminated constant force, so it has to be either an increasing force, a decreasing force, or no force. Those are the only other possibilities.

**Teacher:** When you say "force," we are talking about the extra or "net" leftover force, right?

**Student:** Yeah. Those are the only possibilities for constant velocity. Either the net force is constant, which we eliminated by doing the experiment, or the net, extra force is decreasing, increasing, or zero.

**Student:** Well we know it can't be increasing, 'cuz that just wouldn't make sense that a constant force gives constant acceleration and an increasing force would give constant velocity.

**Teacher:** Is that right, can we eliminate increasing net force then? (Most students agree, even though this was the idea that most believed true when we started.)

**Student:** So it's gotta be a decreasing extra force. (Other students seem to be agreeing, and I want to be sure they don't just accept this idea without its being challenged.)

**Teacher:** Does that make sense that it could be a decreasing net force? (This seems to provide an opening for other students to challenge that hypothesis.)

**Student:** No, that won't work. (Pause) Suppose at first we have a big net force. The cart would have a big acceleration, right? Then, if net

force decreases, we would have a smaller acceleration. Then if net force decreases even more, we have an even smaller acceleration. But we always have an acceleration. It is just getting smaller and smaller, but it is still acceleration.

**Student:** Oh, yeah. In our experiments when there was a big pulling force, we got a big constant acceleration, and with a small pulling force we got a small acceleration.

**Student:** So it has to be zero net force to explain constant speed in a line, logically.

**Teacher:** And why do you say “logically”? (I want all the students to follow and reflect on the argument.)

**Student:** The only possibilities are increasing, decreasing, staying the same or zero. If we eliminated the first three, zero net force is all that is left.

**Student:** But if it was zero net force, how would the cart even get going?

**Student:** In the beginning, the net force would not be zero: that would not accelerate the cart. But once it got going, no net force would be needed. You might need a little bit of force to match any friction, but there wouldn't need to be any extra, leftover force bigger than friction.

**Student:** Wow. That would be like in space. You would just push a little to get the thing going, and then it would go on its own out through space with no force acting on it.

**Teacher:** Provided you didn't need to worry about any gravitational forces.

OK, summarize the results from your experiments. What are your conclusions about the relation between force and motion of various kinds. Also, go back to what you wrote in your journals about force and motion. Are there any situations that you cannot explain with the ideas we developed in class in the last couple of days?

It turns out that these ideas we came up with are consistent with those expressed by Sir Isaac Newton. He wrestled with some of the same concerns you all have expressed before he came to formulate Newton's Laws of Motion. Good job. Good thinking. Tomorrow I want also to summarize the thinking processes by which we came to our conclusions.

The next day I will get the students to summarize the conclusions and to share how the ideas apply to the various situations they had written. They need to see that the new ideas are useful in everyday situations as well as in the class laboratory.

For the last few days the agenda has primarily concerned coming to understand concepts of force and motion. We could not do this without also addressing some aspects of scientific reasoning processes and skills. In this next discussion I want the ideas about the nature and processes of science to be the main question. I want students to be able to answer questions relating to how we come to know in science. Prior units of work—measurement and kinematics—were primarily descriptive. This unit on dynamics and several subsequent units, all of them involving making sensible explanations for phenomena, should offer a promising start to understanding the processes and skills of scientific inquiry.

### Coming to Know and Understand Through Experiment

**Teacher:** We have reviewed the explanations for constant accelerated motion and for constant speed in a straight line. We also have seen how those explanations can be applied to make sense of everyday situations. Now I want us to review how we came to know and believe these explanations for constant acceleration and constant speed in a straight line. What were our thinking processes in coming to those conclusions that might tell us something about doing science?

We now know that constant unbalanced force is needed for constant acceleration and no net force is needed to keep an object going at a constant speed in a straight line, once the object is moving. How did we come to know these ideas? Did you need some authority like the teacher or textbook to tell you the ideas were true? What did we do in class that brought you to believe these ideas?

**Student:** We did experiments.

**Teacher:** But how did we decide what experiments to do?

**Student:** We had some ideas to experiment.

**Student:** The experiments were to test the ideas.

**Student:** Test the hypotheses.

**Teacher:** But where did the ideas or hypotheses come from?

**Student:** You gave us some questions.

**Student:** Those questions were just to get us thinking.

**Student:** Well, but the questions also made us know what our ideas were.

**Student:** We put up the possible ideas we believed at the time, and then the experiment was to test the ideas we thought were true.

**Teacher:** OK, I hear you saying the original questions were important because they helped you identify what your ideas or hypotheses were for explaining the situations. Were your original ideas right? Did the hypotheses have to be right before you could do the experiment?

**Student:** We thought they were right. But they turned out to be wrong.

**Student:** The hypotheses were just our ideas we thought were true before we did the experiment.

**Teacher:** Scientists wonder about things. They might come up with their own questions or they start thinking about questions somebody else has asked. They think they have some ideas to explain what happens. The scientists treat these hypotheses as tentative: maybe they are true, maybe not. The ideas have to be tested. Does this sound like what happened in our class? (Note this was a mini-lecture, which in this case ended with a rhetorical question.) Take a couple of minutes to summarize in your learning journal:

1. What was the role of questions in this investigation?
2. What was the role of the hypotheses?
3. What were the variables involved in our hypotheses?

(After about two minutes the discussion continues.)

**Student:** I think I'm OK with the first two questions, but what were the variables involved?

**Student:** The motion of the cart and the extra force.

**Student:** Oh, yeah. The spring scale reading for force and the dots for motion.

**Teacher:** OK? Any other questions there? (Wait-time.)

In addition to questions and hypotheses, we were talking about experiments. What came out of the experiments? What was the value of doing them?

**Student:** It let us test our hypotheses, to find out whether they were right.

**Teacher:** Say more. How did the experiment do that, specifically with respect to what you did in the experiment?

**Student:** Well, we ran the carts across the table and we graphed the motion.

**Student:** And we read the spring scale while the cart was moving.

**Teacher:** And what did you see? What did you observe? (Note that here the questioning is more of a quick give and take, because we are summarizing the argument, our observations, and our inferences.)

**Student:** That the dots got farther and farther apart.

**Teacher:** Good, so what did that tell us?

**Student:** That the motion of the cart was a constant acceleration.

**Teacher:** That the motion of this cart was constant acceleration. And what did we learn from observing the spring scale?

**Student:** It stayed constant, after it dropped. Why did it drop in the beginning? I'm still not sure about that.

**Teacher:** That is a good question, and we should be able to answer that, but let's stay with this argument for now, OK? (I want to encourage students to ask their own questions, but I want them not to lose sight of the path of this discussion.) OK, so the pulling force stayed constant while this cart was moving across the table.

The constant acceleration and the constant force scale reading were our results from this particular run of the cart going across the table. These were our results of the particular experiment. What happened in the rest of the experiments done by people at other tables and using different masses hanging on the spring?

**Student:** We all got constant acceleration and constant spring force.

**Teacher:** Is that true? (Most students chime in, in agreement.)

**Teacher:** Good. Now notice that you each did your experiment, and you specifically may have gotten a pulling force of 0.8 newtons and you may have gotten 0.7 m/s/s for the acceleration. These are the specific results of the one experiment. But others did similar experiments and you did similar experiments with other specific results, but when we look across all the experiments, we are able to make some tentative conclusions about force and motion. Our conclusions are that when we have any object undergoing a constant acceleration, a constant unbalanced force will be needed to explain the motion. Results are our specific, observable outcomes (or nearly observable, if we have to do a

little bit of work to get the acceleration numbers) from the experiment. Conclusions are our inferences we make to try to summarize the meaning from all those results. Once we have the general conclusions, we try to apply them to lots of specific situations. If the conclusions seem to work, we think we have learned something about the world.

**Teacher:** Now take a couple of minutes to summarize in your learning journals:

4. What is the role of experiments in doing science?
5. What is the difference between results of experiments and conclusions from experiments.

After about three minutes I asked the students to summarize the various answers they had written for each question 1 through 5, since these were the ideas and processes that were the focus of this discussion. Next we needed a short discussion to consider the logical argument used in developing an understanding of forces on the object moving in a straight line with constant speed.

### Coming to Know and Understand Through Logical Reasoning

**Teacher:** Let's think a bit about how we came to believe that straight line motion with a constant speed could be explained by a zero net force. We couldn't do a direct experiment here, but Chris suggested a "logical" argument. Can somebody summarize that for us?

**Student:** Well there were only certain possibilities. Either a constant net force, a decreasing net force, an increasing net force, or no net force was possible.

**Student:** No other outcomes were possible.

**Teacher:** Then what?

**Student:** We eliminated each of the other possibilities, so the only one left was zero net force.

**Teacher:** So only "zero net force to explain constant velocity" was left. Did we verify this by experiment?

**Student:** Not exactly. All our experiments showed if we didn't have zero net force, we got acceleration.

**Teacher:** OK, good. Notice then that sometimes we need to use logical reasoning to determine all the possible explanations and then

eliminate as many as we can on the basis of experiment. So, sometimes experiments tell us more about what the explanation CAN'T be than about what is the right explanation.

Now we have some ways of explaining constant accelerated motion and constant speed in a straight line. We believe these ideas because our experiments together with logical argument seem to support them. Be sure you have these ideas summarized in your learning journal. We will have lots more opportunities to practice these and other processes of scientific inquiry.

## SUMMARY OF LESSONS APPLIED TO THE VIGNETTE

### Setting Learning Goals

In the previous dialogue, my goals included both development of conceptual understanding and development of skills for conducting inquiry. To do that I put the conceptual understanding in the foreground of the activities initially and the inquiry skills, while important, were in the background. In the later discussion, I put in the foreground the development of inquiry skills and the conceptual understanding, while still important, was in the background. This is my way of addressing one of the concerns brought out in other chapters of this book. I cannot teach content without reasoning processes, and I cannot teach reasoning processes without content. But I need to bring explicitly into the foreground whichever I want students to attend to at a particular time. I cannot just teach for development of subject matter using processes and expect students to grasp the skills without setting class time when I ask students explicitly to pay attention to them. If I had just taught for conceptual development, students would not likely have learned inquiry skills, even though I believed I was teaching by inquiry. Learners need to have their attention concentrated on the goals of the learning.

These goals for content and process are included in the national learning goals and in my state guidelines, which suggests that they are for all students. I have taught these ideas to high school physics students and to special education classes of younger students in a physical science course. The latter group requires more time and more carefully centered attention, but nearly all students at upper middle and high school are capable of learning these important ideas and skills.

The laboratory experiences and discussion would not have made sense without some prerequisite knowledge. We had studied the topics of measurement that allowed us to have some idea of how certain we could be with respect to the

measurements we made. We developed skills in the use of measurement apparatus. Our prior study of the description of motion allowed us quickly to analyze the motion and recognize constant acceleration when it occurred. Immediately prior to this investigation of forces in moving situations, we studied forces on objects in the static situations. That sub-unit included using arrows to represent forces and some beginning ideas about the nature of force as an idea used to explain situations. These and other earlier learning experiences helped prepare the way to develop understanding of these difficult ideas about force and motion.

This was not a full, open-ended investigation. I elected to guide it. "Guided inquiry" is probably an accurate description of the process. The ideas to be developed are subtle and not likely to be invented and justified without guidance. These are ideas, moreover, on which much of the rest of our development of content will rest. Yet the investigation was still centered in the students in the sense that their ideas were being tested and the final explanations were constructed by the students from their observations and inferences. This sort of guided inquiry has occupied much of the activity of my classes.

Sometimes I do have the students do open inquiry. But it is typically about ideas that are not going to be so important to our learning story line. For example, I frequently set an open-ended investigation around factors that affect friction: "Identify factors you believe might affect frictional force. Choose two of those factors and design experiments to see whether each factor affects friction, and if so, how each factor is related (mathematically) to friction." Here the primary concern is for inquiry skills. Conceptual understanding of friction is secondary. If I let students do the full investigation with no guidance, they are not likely to come to the idealized conclusion of the science text.

## Dialogues

While questions of fact, such as asking for an observation, can be asked and answered rather quickly, questions for which thinking is required to develop an answer take time and need environments free of distractions. In this vignette, questions from the teacher were used to stimulate thinking on the part of the participants. Many of the important questions to foster development are higher level questions, requiring substantial inferencing.

Dialogues between teacher and students, between student and student, and even within the student are all important. In the preceding discussion, the teacher modeled appropriate interaction between learners by listening to students' ideas and responding by respectfully asking questions and gently challenging ideas and

reasoning that didn't yet make sense. The intention was to stimulate reflection by the students. Students were encouraged to ask and attempt to answer one another's concerns. Wait time and other appropriate listening techniques were used to foster personal reflection by students. It was hoped that eventually students would internalize the dialogue and probe and stimulate their own thinking.

Notice that there were virtually no evaluative comments by the teacher. Typical verbal interaction in a classroom involves the teacher's asking a question, the student responding, and the teacher's evaluating the response. What happens in my class is more like a conversation among the several participants. I attempt to build an atmosphere of collaboration among participants who are collectively constructing meaning through collective inquiry. This classroom climate requires mutual respect among the participants. Students need to support one another as they share their ideas and gently probe for understanding. Students need to trust the guidance of the teacher: to believe that the teacher is putting experiences before them that will allow them to make sense. Teachers need to respect learners in their attempts to make sense, respecting that they are capable of the skills and reasoning that will allow them to develop desired understanding.

### From Experiences to Ideas

In this vignette, students were not told Newton's Laws at the outset. They began with typical experiences of books sliding across tables. They suggested hypotheses about motion that made sense to them at the outset. After experiences with the experiments they reconsidered their initial ideas and suggested new understandings about force and motion. Then I confirmed that the conclusions they had formed were consistent with Newton's Laws. The experiences came first, then the idea, and finally the confirmation from authority. Following this line of development, students suggest, makes them feel empowered, capable of coming to understanding through their own resources and capabilities. They don't need some authority to tell them what is true about the natural world. I did conduct short lectures, but they were after students had sufficient experiences with hand and mind that what I had to say would make sense.

### Materials to Support Inquiry

Equipment and materials to support inquiry do not need to be extensive. The activities in the vignette consisted of a handout of elicitation questions to which students were to respond and the simple apparatus involving cart, spring scale,

string and pulleys. The tickertape timer was more sophisticated, but students had enjoyed an opportunity earlier to learn how to use that to create a motion story. How the equipment worked was relatively transparent. More important is that the equipment was just enough to address the goals for learning, to measure the relevant variables from which students could test their initial ideas.

After students had developed the ideas consistent with Newton's Laws, text materials were made available to students. My course was not directed by a textbook. At its beginning, students had three texts from which to choose for background reading. That not all students were checking out the same text helped to emphasize the freedom of the course from the constraints of a textbook. I did use technology that supported the inquiry students had experienced. After guiding their inquiry on the development of the ideas for explaining different sorts of motion, I asked them to interact with a computerized Diagnoser to check whether their knowledge and reasoning were consistent with the results of our inquiry. Diagnoser serves questions in pairs. Following a phenomenological question essentially asking "What would happen if...?" is a question asking "What reasoning best supports the answer you chose?" Expected responses were associated with one or another's attempt at understanding that falls short of the learning target. Immediate feedback for such responses suggested what students need to think about, how what they said or did was inconsistent with experiences from classroom activities or from daily life (Hunt & Minstrell, 1994, and see [www.talariainc.com/facet](http://www.talariainc.com/facet)). To support the inquiry, the computerized Diagnoser works to build ideas from experiences and in a manner consistent with dialogues.

### Final Comment

Inquiry teaching is hard work. Students may ask questions for which I don't know the answer. For me this has been stimulating and a great source of learning. It would be much simpler to lecture, but I wouldn't learn as much myself. I have to learn the various directions students might go. I need to make decisions about which of those directions I am willing to invest the intellectual energy of the class in investigating.

Inquiry teaching takes time. It would be faster to tell the students the "correct" scientific ideas. But to learn to do inquiry, students need opportunities to participate in learning by inquiry. The fruits of the additional time investment include deeper and lasting understanding. It also fosters increased opportunities for life-long learning for both students and teacher.

## REFERENCES

- Haber-Shaim, U., J. Cross, J. Dodge, and J. Walter. 1971. *PSSC physics*. Lexington MA: D.C. Heath.
- Hunt, E., and J. Minstrell. 1994. A cognitive approach to the teaching of physics. In *Classroom lessons: Integrating cognitive theory and classroom practice*, edited by K. McGilly, 51-74. Cambridge, MA: MIT Press.
- Minstrell, J. 1982. Explaining the “at rest” condition of an object. *The Physics Teacher* 20: 10-14.
- Minstrell, J. 1984. Teaching for the development of understanding of ideas: Forces on moving objects. In *AETS Yearbook: Observing Science Classrooms*, edited by C. Anderson, 53-72. Columbus OH: Ohio State University. Association of Educators of Teachers of Science.
- Minstrell, J. 1992. Facets of students’ knowledge and relevant instruction. In *Proceedings of the international workshop: Research in physics learning—theoretical issues and empirical studies*, edited by R. Duit, F. Goldberg, and H. Niedderer, 110-128. Kiel, Germany: The Institute for Science Education at the University of Kiel (IPN).
- Rowe, M.B. 1974. Wait-time and rewards as instructional variables, their influence on language, logic, and fate control: Part one—wait time. *Journal of Research in Science Teaching* 11:81-94.
- Rutherford, J., H. Holton, and F. Watson. 1972. *Project physics*. New York: Holt, Rinehart and Winston, Inc.
- van Zee, E.H. and J. Minstrell 1997. Using questioning to guide student thinking. *The Journal of the Learning Sciences* 6: 227-269.