Assessment and Instruction in the Science Classroom

CSE Technical Report 418

Gail P. Baxter and Anastasia D. Elder
University of Michigan

Robert Glaser
CRESST/Learning Research and Development Center
University of Pittsburgh

November 1996

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported under the Educational Research and Development Center Program, cooperative agreement number R117G10027 and CFDA catalog number 84.117G, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

ASSESSMENT AND INSTRUCTION IN THE SCIENCE CLASSROOM

Gail P. Baxter and Anastasia D. Elder
University of Michigan

Robert Glaser
CRESST/Learning Research and Development Center
University of Pittsburgh

Abstract

Changes in knowledge underlie the cognitive capabilities that are displayed in competent performance and the acquisition of improved performance. It is important to bring these knowledge-generated processes to attention because they represent possibilities for instructional design that might improve learning. In this paper, the role of performance assessments in making relevant cognitive activity apparent to teachers and students is discussed. Descriptions of the cognitive activity of fifth-grade students while carrying out a science performance assessment reveal critical differences between those who think and reason well with their knowledge of circuits and those who do not. Differences in quality of explanations, adequacy of problem representation, appropriateness of solution strategies, and frequency and flexibility of self-monitoring indicate more or less effective learning of the subject matter. Awareness of and attention to these cognitive characteristics of competent performance in an assessment situation provide teachers the necessary feedback to construct classroom environments that encourage reasoning and knowledge integration. In this way, performance assessments not only evaluate student performance but suggest changes in instructional practice to support effective learning in the elementary science classroom.

Assessing Knowledge-Based Competence

Studies of human cognition have made strong contributions to understanding how individuals construct and structure their knowledge as they become increasingly skilled and competent in subject matters they learn in and out of school (cf. Bereiter & Scardamalia, 1987; Charles & Silver, 1988; Chase & Simon, 1973; Gobbo & Chi, 1986; Schoenfeld, 1992). Much of this work has compared the differences between people who are competent in solving problems and performing complex tasks and beginners who are less proficient. Results of numerous studies suggest that as a result of learning, children and adults develop special features of their knowledge that contribute to their ability to think and reason with what they know. Further, declarative knowledge is integrated with an understanding of when and how to use that knowledge. The resultant knowledge structure enables certain cognitive activities such as generating and elaborating explanations, building a mental model or representation of a problem to guide solution, managing thinking during problem solving to efficiently allocate resources, and enlisting appropriate, goal-directed solution strategies to facilitate problem solving (Glaser, 1991). Performance capabilities such as these are indicative of effective learning and experience with a body of knowledge and may be appropriately used to define achievement in a subject matter domain (Glaser & Silver, 1994).

The conceptualization of student achievement and competence in terms of the quality of cognition and the ability for thinking has influenced educational practitioners and policy makers as they turn to performance assessments as a major instrument of reform. An underlying belief motivating these efforts is that these assessments, if based on modern knowledge of cognition, would provide models and standards of practice for students and teachers. It is this relationship between the cognitive activity involved in assessment and teaching practice that provides the context for this paper. Broadly speaking, our purpose is to demonstrate ways in which assessments aligned with instruction and theories of knowledge development can help teachers and students attend to the relevant cognitive activities underlying knowledge-based performance. Making the thinking of the learner overt (put on display so to speak) provides opportunities for it to be examined, questioned, and realized as an active object of constructive teaching and the focal point of assessments. Further, linking assessments and learning with the processes of competence provides teachers the necessary feedback to construct classroom environments that encourage reasoning and knowledge integration.

Cognitive Expectations

The nature and quality of cognitive activity enabled by different levels of conceptual and procedural knowledge are suggested by comparative analysis of experts and novices in various domains (cf. Chi, Glaser, & Farr, 1988). These studies and others provide the conceptual and empirical basis for our examination of the reasoning and problem-solving activities that support inferences of knowledge development in the science classroom. In brief, competent students

(a) provide coherent explanations based on underlying principles rather than descriptions of superficial features or single statements of fact; (b) generate a plan for solution that is guided by an adequate representation of the problem situation and possible procedures and outcomes; (c) implement solution strategies that reflect relevant goals and subgoals; and (d) monitor their actions and flexibly adjust their approach based on performance feedback (see Table 1).

Using these critical characteristics as a framework, the kind and quality of cognition that science performance assessments demand of students were examined in a series of studies (e.g., Baxter, Elder, & Glaser, 1994; Baxter, Glaser, & Raghavan, 1993). Student protocols and observations were collected and analyzed for tasks that required students to reason with subject matter knowledge to solve problems (e.g., Baxter, Elder, & Shavelson, 1995; Shavelson, Baxter, & Pine, 1991), engage in an extended inquiry of an everyday phenomenon (e.g., Baron, Carlyon, Greig, & Lomask, 1992), or combine their understandings of physical, life, and earth science to make decisions in real-world contexts (California State Department of Education, 1993). For our purposes here, we discuss the expected and observed cognitive activity of students while carrying out the Electric Mysteries assessment. The primary intent is to demonstrate how the nature and extent of cognitive activity underlying task performance permit inferences about student understanding and subject matter knowledge.

In the following sections, we provide examples of how the cognitive activity of students on a fifth-grade science performance assessment can highlight opportunities for instruction to foster thinking and reasoning with acquired

Table 1

Quality of Cognitive Activity Enabled by Level of Knowledge

	Level of knowledge	
Cognitive activity	Low →	High
Explanation	Single statement of fact or description of superficial factors	Principled, coherent
Plan	Single hypothesis	Procedures and outcomes
Strategy	Trial-and-error	Efficient, informative, goal oriented
Monitoring	Minimal and sporadic	Frequent and flexible

knowledge. First, we briefly describe the circuits instructional unit and the Electric Mysteries performance assessment. Next, we outline the cognitive expectations for this assessment derived from modern theory of developing competence. Then, we provide detailed descriptions of student thinking and reasoning as concrete examples of varying competence in the classroom. The paper concludes with suggestions for making thinking overt in the science classroom and a discussion of the importance of the cognitive components of competence as a framework for teaching, learning, and assessment.

Assessing Competence in the Classroom: An Example

As science education moves to embrace the constructivist notions of teaching and learning, hands-on programs are seeing renewed popularity in elementary classrooms. This change in instructional practice has been accompanied by efforts to develop alternative forms of assessment consistent with these views of teaching and learning. Consider an instructional unit on electric circuits. Teachers guide students through a series of activities intended to foster their understanding of the nature of a circuit and the effect of changing various components in that circuit. Students working in groups generate hypotheses, debate alternative solution strategies, and draw conclusions on the basis of their investigations.

To assess the extent of student learning at the end of the instructional unit students are provided with some equipment and asked to reason with their subject matter knowledge to identify the circuit components in each of six boxes (e.g., Shavelson & Baxter, 1992). Specifically, students are asked to determine the contents of each of six “mystery” boxes, A through F, from a list of five possible alternatives (see Figure 1). Students are provided with two batteries, two bulbs, and five wires to construct circuits to test each of the six boxes. Two of the boxes contain the same thing (boxes B and F each have a wire). All of the others have something different (battery and bulb, two batteries, nothing, or a bulb).

In carrying out the assessment, students engage in a cyclical process of hypothesis testing and refining to identify the circuit components enclosed in each of the six boxes. Using their knowledge of what constitutes a circuit, and the impact of changing various components in a circuit (e.g., adding a second bulb), students test their hypothesis by observing a bulb connected in a circuit external to the box. For example, if the bulb is dim when connected in a circuit to one of the

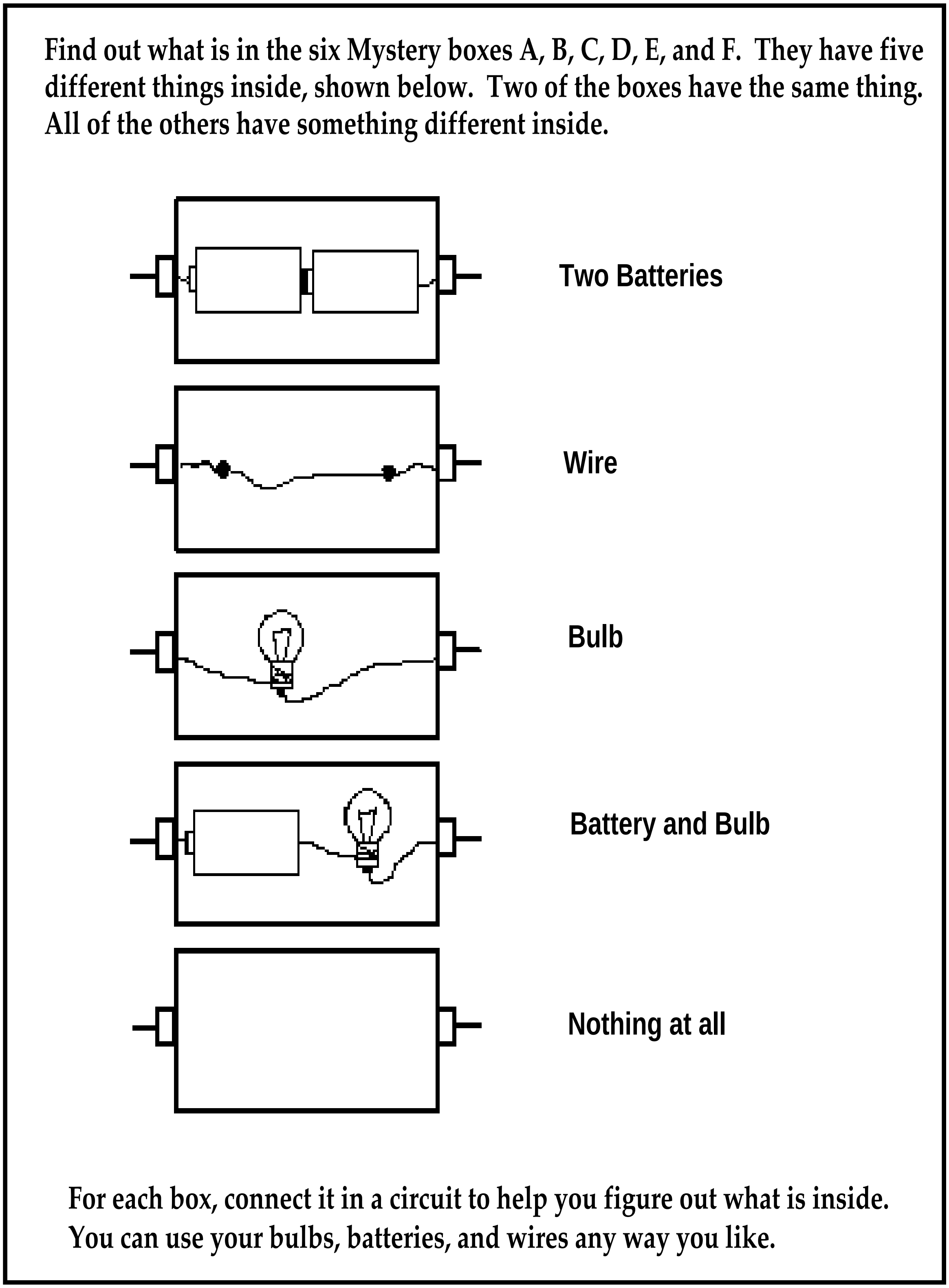

Figure 1. Electric Mysteries assessment (adapted from Shavelson, Baxter, & Pine, 1991).

boxes, students might reason that a battery and a bulb are in the box. If the bulb is very bright, students might reason that there are two batteries in the box.

As part of a larger study, this Electric Mysteries assessment was administered individually to 31 students (15 females and 16 males) enrolled in a hands-on science program in an urban school district in southern California.[1] At the time the assessment was administered, students had recently completed an eight-week unit of study on circuits. Prior to carrying out the assessment, instructions were read to the students and the equipment was introduced. Students were told that the interviewer was interested in what they and other students think when they do science, and that they needed to talk out loud while carrying out the investigation so their thinking would be apparent (cf. Ericsson & Simon, 1993). Using the characteristics of competent performance described above, we gathered information about students' cognitive activity (i.e., students' explanations, plans, strategies, and monitoring) while carrying out the Electric Mysteries performance assessment.

Explanation. Effective learning of content knowledge enables students to explain principles underlying their performance (e.g., Chi, Bassock, Lewis, Reimann, & Glaser, 1989; Fay, 1995). After being oriented to the equipment, students were asked to construct a circuit using one bulb, and one battery and wires. Then they were asked, "Can you tell me what a circuit is? Can you tell me how a circuit works?" as a way to elicit information about students' task-specific conceptual understanding. Accurate, coherent, and complete explanations of a circuit and how it works suggest students have a developed knowledge of circuits. Fragmented explanations suggest a limited knowledge of circuits.

Plan. Competent individuals qualitatively assess the nature of a problem and construct a mental model or internal representation prior to initiating a solution strategy (cf. Gentner & Stevens, 1983; Halford, 1993). This representation is used to anticipate alternative outcomes to various actions (e.g., the bulb does not light when connected in circuit with a box) and to generate next steps based on those outcomes (e.g., test with a bulb and battery in the circuit).

Before starting the investigation, students were asked: "How are you going to go about solving this problem? That is, how will you determine what is in each of the six mystery boxes?" Plans composed of actions and anticipated outcomes, a

---

[1] For details of the study, the reader is referred to Baxter, Glaser, and Raghavan (1993).

sort of trial run through the solution strategy, suggest well-developed knowledge of circuits. Lack of planning prior to manipulating the equipment suggests ineffective learning.

Strategy. Principled problem solving is characterized by the use of goal directed, efficient strategies and is reflective of substantial knowledge organization and structure (e.g., Siegler, 1988). Constructing test circuits in a systematic and purposeful fashion—connect each box with a bulb to determine which of the six boxes has a battery, and then connect the remaining boxes with a battery and bulb—suggests a well-developed knowledge of circuits. A seemingly random sequence of circuit construction in a trial-and-error fashion suggests ineffective learning.

Monitoring. Frequent, flexible monitoring is a hallmark of competence (Glaser, in press). In carrying out the Electric Mysteries assessment, students should attend to and coordinate knowledge of circuits, knowledge of task constraints, and interpretations of current trials. Simultaneous attention to these pieces of information demands that students apply a range of monitoring skills to check their thinking and reasoning throughout their investigation.

Four types of monitoring activity were identified for the Electric Mysteries assessment. They are (a) Circuit Comparison: Student compared bulb brightness between boxes or between a box and an external circuit; (b) Hypothesis Retesting: Student retested boxes to check his or her results; (c) Constraint Checking: Student referred to and accounted for the possible contents of the boxes provided in the task instructions; and (d) Problem Recognition: Student acknowledged an inconsistency between his or her observations and hypothesis. Performance characterized by engagement in a variety of monitoring activities in a manner consistent with the demands of the task suggests a substantial knowledge of circuits. In contrast, lack of attention to performance feedback suggests an inadequate knowledge of circuits.

Exemplars of Student Cognition

An examination of student protocols suggested three qualitatively different patterns of cognitive activity: consistently high, intermediate, and consistently low. The consistently high and consistently low groups perform much like the descriptions of high- and low-knowledge students in Table 1. That is, their performance across all four cognitive activities was consistent with high- or low-knowledge students. Students in the intermediate group displayed cognitive activity indicative of *some* knowledge of circuits; this knowledge, however, could not support the quality of performance displayed by students in the consistently high group. Next, we provide a general summary of the performance of students in each of these three groups accompanied by an in-depth description of one student who typifies the thinking and reasoning of students in that particular group.

Consistently High Levels of Cognitive Activity

Twenty-three percent of the students demonstrated consistently high levels of cognitive activity. These students (a) provided a clear, correct explanation of a circuit; (b) articulated a plan that anticipated the outcomes of various strategies; (c) displayed a systematic approach to solving the problem by testing first with a bulb in circuit and then, if necessary, testing with a battery and bulb in circuit; and (d) engaged frequently in a variety of monitoring activities (e.g., used external circuit as a standard for comparison of relative bulb brightness, referred to instructions for list of options) to check their thinking and reasoning and adjust their performance as necessary.

Carlos' performance is typical of the students in this group. When asked how a circuit works, Carlos provided an explanation that incorporated the notion that electricity flows in a circular pathway within a closed system. He responded,

> It works by just electricity flowing. I mean electricity flowing through the wires connecting there to the light bulb and this connects to the minus part of the other battery and with that you have both sides full so that you can receive both electric currents and you come back to the bulb . . . It goes through there and comes back around like that.

In generating a plan, Carlos had a goal in mind and some strategies to reach that goal. He stated, "I'll probably start with a battery, a light bulb, and maybe two wires and put them everywhere and that way . . . if the light just shines regular, that will be [the wire] and if it shines really bright, that will be [two batteries] probably, and if it doesn't shine at all, it will be [nothing]." His statement suggests that in planning his investigation, he relied on his knowledge of circuits to mentally run through a solution strategy, the potential outcomes of that strategy, and interpretations of those outcomes in terms of the problem he was trying to solve.

As described in his plan, Carlos systematically connected each box to a battery while carrying out the assessment. Because a battery and bulb in circuit with a box did not always provide the confirming evidence needed to reach a conclusion with some confidence, Carlos adopted other strategies. The strategy he chose depended on the outcome of testing with a battery and bulb in circuit. When appropriate, Carlos connected two batteries and a bulb in circuit, or constructed an external circuit (one without the box attached), or reversed the direction of the battery in circuit to account for polarity.

The most efficient procedure—test each box first with a bulb and then with a battery and bulb—was not characteristic of Carlos' performance. Nevertheless, what Carlos lacked in efficiency, he compensated for with his knowledge of circuits. For example, when he identified more than one pair of boxes as having the same contents, he considered several possible interpretations for his results: "Maybe one of the batteries is dead or minus is connected to minus. You got to have negative connected to positive or positive connected to negative." By reversing the direction of the battery in circuit with the box, he could rule out one or more options from the list of possible contents listed in the assessment instructions (see Figure 1).

Carlos continually monitored his performance as he conducted his investigation of the "mystery" boxes. In all he displayed 17 instances of monitoring, including:

a. Circuit Comparison: "It must be the same thing as A. . . . Because it is uh, because like this is, you see, the exact same thing as the other one was, I could tell that by the way I experimented, I saw how bright it was like this."

b. Hypothesis Retesting: "Now, I am going to go back to A and B because they might be different."

c. Problem Recognition: "They might be, they might be different ones and maybe I could discover that because I didn't notice it before because maybe E, maybe E and C are the same, I don't know . . ."

The frequency and flexibility of monitoring displayed by Carlos helped him successfully operate within the constraints of the task. Recall there were only five possible circuit components enclosed in the six boxes; two boxes had the same circuit components inside. To determine the contents of each box requires simultaneous attention to the relative brightness of the bulb in circuit with each of the boxes *and* attention to one's conclusions *and* attention to the list of options. When Carlos was uncertain about the absolute brightness level of the bulb connected in circuit to a box, he created an external circuit that served as a standard to judge the relative brightness. Further, when he noted he had used an option more than once (i.e., constraint checking), he retested boxes to reassess his conclusions.

Intermediate Levels of Cognitive Activity

The quality of cognitive performance displayed by 60% of the students is best described as *intermediate*; these students did not display consistently high or consistently low levels of cognitive activity. Although explanations from these students were partially correct (circuit is a closed system) they often included some alternative conceptions of how a circuit works (e.g., explosion model). Students in this group could not generate an adequate representation of the task that would facilitate their thinking through a possible solution strategy in advance; their representations were generally restricted to the impact of adding a bulb in circuit to the boxes. Their problem-solving strategies were inefficient and unsystematic; they tended to repeat circuits or try many different circuits. Although most of the test circuits were potentially informative (i.e., bulb, or battery and bulb), these students could not always interpret the outcomes of their tests. Finally, students in this group did monitor their performance, albeit sporadically (i.e., lacking in frequency and flexibility).

Dana exemplifies students in this group. Her explanation incorporated the notion of a complete pathway similar to Carlos' above: "You have to have a full circuit to make the bulb light because if you don't have a full circuit, you won't have ah, a pathway to make the bulb light." However, she went on to state,

> . . . usually the energy is going through the wires and then the metal is taking it because you have it hooked here and the metal is taking it there and the light is touching the metal so it, the energy is going through all of that and it is lighting.

Dana attempted to clarify her explanation by showing the interviewer what she meant; she gestured that energy from the battery travels along two (seemingly independent) routes terminating in an explosion inside the bulb. Dana's demonstration and explanation implied that she was thinking about electrical flow as analogous to a meeting of two one-way streets. Although she could identify the necessary components in a circuit and articulate that a circuit involves a pathway, Dana had little understanding of how electricity flows in a circuit.

In generating a plan Dana stated, "First, I am going to try the wires and the bulbs in all of them [the boxes] to see if it works and if they light, and if it doesn't, then I'll try the battery and the wires." Her plan was not elaborate in the way that Carlos' plan was; she did not link particular kinds of circuits (e.g., bulb) to specific outcomes (e.g., dim) and interpretations of those outcomes (e.g., battery and bulb is in the box). Further, she failed to anticipate the ineffectiveness of connecting a battery in circuit to the box. Rather, her plan consisted of naming the equipment and the sequence in which she would use it.

Although at first glance, Dana's plan might appear to be somewhat haphazard, it is reflective of her knowledge of circuits. That is, Dana knew that a circuit requires, at a minimum, a battery, bulb, and wires. In thinking about her approach, she indicated that she will connect the "missing" component to complete the circuit. If she thought the box contained a bulb, she would connect a battery and wires in circuit; if the box contained a battery, she would connect a bulb and wires in circuit; if the box contained a wire, she would connect a battery and bulb in circuit. Dana's limited understanding of circuits and the impact of changing various components in the circuit constrained her ability to anticipate the outcomes of her proposed plan.

In carrying out her investigation, Dana for the most part followed the plan she articulated. Her strategy consisted of testing each box by first connecting it in circuit with a bulb, and then connecting it in circuit with a battery. She consistently maintained this strategy on all of the boxes except box D. (The bulb lit brightly when connected in circuit to box D because box D contained two batteries.) She believed that connecting a single battery was informative in her attempt to identify the components in the boxes because the warmth of the wire indicated energy passing through. She explained, "I am feeling it to see if it is getting warm, if energy is going through it . . . Energy is going through it if there is a light. I think there is a light in this one." And later, on another box, she remarked, "I am feeling it to see if the wires are getting hot and this one I think that there is nothing in there because the wires are not getting hot." After testing all the boxes, Dana reviewed her answers and realized that she had not found any boxes that contained a wire. Given her understanding that a circuit is constructed with three components—battery, bulb, wires—she decided to connect a battery and bulb in circuit to test for a wire inside. Apparently, she did not recognize the utility of systematically testing all boxes with a bulb and then testing with a battery and bulb in circuit. Nor did Dana have a generalized strategy for solving the task; her choice of strategy was dependent on the goal she set for herself. For example, if she hypothesized that one of the three necessary circuit components was in the box (bulb), then she would connect the remaining components to the box (battery and wires) to complete the circuit.

Dana monitored less frequently (a total of 9 instances) and less effectively than Carlos. She had a variety of monitoring strategies (circuit comparison, hypothesis retesting, and constraint checking) but consistently relied on retesting her hypotheses. Indeed, when she did use another potentially effective strategy such as comparing circuits, she either misinterpreted the evidence or didn't use relevant criteria (i.e., bulb brightness) when making her comparisons. She stated: "I think that there is nothing in there because the wires are not getting hot. And the other one, the wires got warm pretty soon, so I think that the box either [sic] has nothing in it."

After considering her answers (A and D have two batteries, B and F have a bulb, C and E have nothing), she commented: "I think one of these has to be a wire, because there are two of, or each one of them has two; there has to be a change here somewhere." This led her to retest all of the boxes to "check" for a battery and bulb or a wire. Nevertheless, the problem representation that guided her solution strategy (a circuit consists of a particular configuration of a battery, bulb, and wires) constrained the procedures she invoked.

Consistently Low Levels of Cognitive Activity

Seventeen percent of the students could not explain what a circuit is or how it works. Their explanations consisted of a single fact about circuits, such as the battery is a source of energy. When asked for a plan, these students provided a hypothesis for the contents of the first box and then began testing it. Many engaged in a trial-and-error strategy—hook something up and see what happens. Others consistently applied one strategy (e.g., connect bulb only in circuit with each box) without regard for its effectiveness. In monitoring their performance, these students relied primarily on their memory of what had happened with other boxes and not on a set of task-related strategies that would provide appropriate feedback.

Consider Raymond. In his explanation of a circuit, Raymond identified the battery as a source of energy. Although he thought that the electricity flowed through the system in a circle, he thought that it made one cycle and then stopped: "Inside the battery it has energy, and the energy goes in the wire and out the wire, then it goes to the light bulb, then it goes to the other wire, and into the battery again . . . It stops right here in the middle." Raymond, like Dana, had learned how to manipulate equipment to construct a circuit, but did not fully understand what a circuit is or how it works.

In contrast to Carlos and Dana, Raymond could not offer a plan to guide his solution. Rather, he generated an hypothesis for the contents of box A. "Um, I think this is a bulb." Then, he immediately began connecting a bulb in circuit to box A. Raymond approached the investigation with one strategy in mind—connect a bulb in circuit to each box. This strategy was effective when there was a battery in the box—box A has a battery and bulb, and box D has two batteries—but ineffective for the other boxes. Nevertheless, Raymond consistently applied this strategy regardless of its adequacy; he did not appear to recognize the limitations of his approach. That is, he did not take advantage of the feedback from the task to change his strategy.

Although Raymond did attempt one form of monitoring—that of comparing the relative bulb brightness in two circuits—he did so only once. He concluded that box D contained two batteries "because, um, the light bulb lights up more than A." This comparison was based on his memory of the brightness of the bulb he connected to box A and not on a simultaneous comparison of the circuit connected to box D with that connected to box A. Simultaneous comparison is the most reliable way to compare relative brightness and identify differences and is characteristic of the kinds of comparisons that Carlos made.

Teaching, Learning, and Assessment in the Classroom

Descriptions of the cognitive activity of fifth-grade students while carrying out the Electric Mysteries science performance assessment reveal critical differences between those who think and reason well with their knowledge of circuits and those who do not. In general, students displayed one of three qualitatively different levels of cognitive activity: consistently high, consistently low, or intermediate. Students displaying consistently high levels of cognitive activity described a plan consisting of procedures and interpretation of possible outcomes, expressed through their explanations an understanding of the conceptual knowledge of circuits, demonstrated an efficient, principled approach to solving the problem, and engaged in frequent and flexible monitoring. In contrast, students displaying consistently low levels of cognitive activity offered a hypothesis when asked for a plan, provided a factual statement when asked for an explanation, invoked a trial-and-error strategy of "hook something up and see what happens" to guide their problem solving, and monitored their performance sporadically at best while carrying out the Electric Mysteries assessment. The performance of a majority of students (characterized as intermediate) demonstrated that students had some understanding of circuits, but their knowledge was not sufficiently structured to sustain high levels of reasoning and thinking throughout the assessment. These students generated plans and explanations that were accurate but incomplete. Their procedural strategies and monitoring were generally informative but insufficient to successfully complete the assessment.

Comparisons such as these demonstrate how students with various levels of knowledge reason on an assessment task and direct attention to the relevant cognitive activity underlying performance. Further, they highlight opportunities for instruction to foster reasoning and thinking with acquired knowledge. Awareness of and attention to these sorts of activities that differentiate more from less proficient performance can support the development of thinking and reasoning in the elementary science classroom. For example, the relatively low performance (80% were intermediate or consistently low) of students in this and other assessment situations suggests a mismatch between instruction and assessment. In the descriptions presented here, it is apparent that a large number of students learned how to manipulate equipment to construct a circuit as part of their science instruction. Their knowledge, however, stopped there. They did not appear to have a well-developed understanding of how a circuit works or an understanding of the effect of changing one or more components in a circuit. This suggests that it is possible to acquire knowledge and skills in ways that preclude thinking.

Because how knowledge is acquired determines how it is used, teachers with performance assessments as tools can design classroom situations in which cognitive activity can be displayed and practiced. Teachers can call attention to the differences between plans that relate to appropriate goals and those that do not, provide students the opportunity to discriminate between effective and ineffective strategies (e.g., eliminating alternatives, breaking down the problem into subgoals), and model self-questions and self-explanations (and other self-monitoring techniques) that regulate the effectiveness of their performance. In addition, they can encourage students to reflect on how their efforts relate to what they are trying to accomplish and on the meaning and relevance of their initial representation of the problem. Strategies for how to represent problems must be taught as well as strategies for how to solve problems. Ultimately, the fruitful integration of teaching, learning, and assessment demands that ". . . these cognitive activities are taught not as subsequent add-ons to what we have learned, but rather are explicitly developed in the process of acquiring the knowledge and skills that we consider the objectives of education" (Glaser, 1984, p. 93). To accomplish this, teachers should understand the cognitive components of effective use of knowledge. Teachers can then appropriately reflect on the ways in which their classroom environment and their own pedagogical practices influence the quality of student cognition. The essential issue is that curriculum-linked performance assessments, based on theories of knowledge development, can make critical cognitive activity and the application of effort relevant and visible to students and teachers. In this way, performance assessments not only evaluate student achievement but also highlight opportunities for learning and instruction.

References

Baron, J. B., Carlyon, E., Greig, J., & Lomask, M. (1992, March). *What do our students know? Assessing students' ability to think and act like scientists through performance assessment*. Paper presented at the annual meeting of the National Science Teachers Association, Boston.

Baxter, G. P., Elder, A. D., & Glaser, R. (1994). *Cognitive analysis of a science performance assessment* (CSE Tech. Rep. No. 398). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.

Baxter, G. P., Elder, A. D., & Shavelson, R. J. (1995). *Effect of embedded assessments on performance in elementary science classrooms*. Manuscript submitted for publication.

Baxter, G. P., Glaser, R., & Raghavan, K. (1993). *Cognitive analysis of selected alternative science assessments* (CSE Tech. Rep. No. 382). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

California State Department of Education. (1993). *Science grade 5 administration manual*. Sacramento, CA: Author.

Charles, R., & Silver, E. (Eds.). (1988). *The teaching and assessing of mathematical problem solving*. Reston, VA: National Council of Teachers of Mathematics.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55-81.

Chi, M. T. H., Bassock, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13,* 145-182.

Chi, M. T. H., Glaser, R., & Farr, M. (Eds.). (1988). *The nature of expertise.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis. Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

Fay, A. L. (1995, March). *Factors affecting the content and structure of children's science explanations*. Paper presented at the biennial meeting of the Society for Research in Child Development, Indianapolis, IN.

Gentner, D., & Stevens, A. L. (Eds.). (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist, 39,* 93-104.

Glaser, R. (1991). Expertise and assessment. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition* (pp. 17-30). Englewood Cliffs, NJ: Prentice Hall.

Glaser, R. (in press). Changing the agency for learning: Acquiring expert performance. In A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games*. Mahwah, NJ: Lawrence Erlbaum Associates.

Glaser, R., & Silver, E. (1994). Assessment, testing, and instruction. In L. Darling-Hammond (Ed.), *Review of research in education* (Vol. 20, pp. 393-419). Washington, DC: American Educational Research Association.

Gobbo, C., & Chi, M. T. H. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development, 1*, 221-237.

Halford, G. S. (1993). *Children's understanding. The development of mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-370). New York: Macmillan.

Shavelson, R. J., & Baxter, G. P. (1992). What we've learned about assessing hands-on science. *Educational Leadership, 49*(8), 20-25.

Shavelson, R. J., Baxter, G. P., & Pine, J. (1991). Performance assessment in science. *Applied Measurement in Education, 4*, 347-362.

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development, 59*, 833-851.