DOCUMENT RESUME

ED 358 160 TM 019 959

| | |
|---|---|
| AUTHOR | Smith, Laura M.; And Others |
| TITLE | Assessment of Student Learning in Science. |
| INSTITUTION | South Carolina Univ., Columbia. South Carolina Center for Excellence in the Assessment of Student Learning. |
| SPONS AGENCY | South Carolina Commission on Higher Education, Columbia. |
| PUB DATE | Mar 93 |
| NOTE | 21p.; For related documents, see TM 019 960-965. Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993). |
| AVAILABLE FROM | South Carolina Center for Excellence in the Assessment of Student Learning, College of Education, University of South Carolina, Columbia, SC 29208. |
| PUB TYPE | Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150) |
| EDRS PRICE | MF01/PC01 Plus Postage. |
| DESCRIPTORS | *Academic Achievement; Computer Simulation; Curriculum Development; *Educational Assessment; Educational Change; Elementary Secondary Education; Evaluation Methods; Performance Tests; *Portfolios (Background Materials); *Science Education; State Programs; *Student Evaluation; Student Journals |
| IDENTIFIERS | Alternative Assessment; Authentic Assessment; Concept Maps; Hands On Experience; Open Ended Questions; *Performance Based Evaluation; Reform Efforts; *Science Achievement; South Carolina |


ABSTRACT

As part of a series of documents describing assessment of student learning in various curriculum areas in South Carolina, this document reviews the assessment of student learning in the science classroom. The report begins with a discussion of effective science learning, and outlines the knowledge, skills, and dispositions that encompass science education. Reform efforts currently underway to transform school science are described. Alternative methods of assessment that support curriculum and instruction and promote exemplary science learning are explored, and a brief summary of activities in science assessment in selected states is provided. Performance assessments that are discussed include: (1) performance tests and tasks; (2) open-ended questions; (3) student journals; and (4) computer simulations. Portfolios and concept maps are other forms of authentic assessment that are discussed. Appendix A presents an example of a hands-on science assessment, and Appendix B presents a problem and a concept map. (SLD)

# ASSESSMENT OF STUDENT LEARNING IN SCIENCE

Laura H. Smith

Joseph M. Ryan

Therese M. Kuhs

South Carolina Center for Excellence
in the Assessment of Student Learning

March 1993

The South Carolina Center for Excellence in the Assessment of Student Learning was established by the South Carolina Commission on Higher Education and the College of Education, University of South Carolina

# CONTENTS

**South Carolina Center for Excellence in the Assessment of Student Learning (CEASL)**

## ASSESSMENT OF STUDENT LEARNING IN SCIENCE

### Foreword

This report is part of a series of documents prepared by the South Carolina Center for Excellence in the Assessment of Student Learning (CEASL) to describe assessment of student learning in various school curriculum areas from prekindergarten through grade twelve. The focus of this document is assessment of student learning in the science classroom. The report begins with a discussion of effective science learning and outlines the knowledge, skills, and dispositions that encompass science education. The next section describes reform efforts currently underway that are designed to transform school science. Alternative methods of assessment that support curriculum and instruction and promote exemplary science learning are explicated in subsequent sections. Lastly, a brief summary of activities in science assessment from selected states is provided.

The South Carolina Center for Excellence in the Assessment of Student Learning was established by the South Carolina Commission on Higher Education and is supported by the South Carolina Commission on Higher Education and the College of Education, University of South Carolina. The purpose of this Center is to increase awareness among teacher-educators of recent efforts to change approaches used to assess students' learning in preschool through high school, and to encourage and support efforts to enhance training in testing, measurement and the assessment of students' learning for preservice educators. The Center is based on the educational philosophy that the fair, accurate and informative assessment of students' learning is an integral part of the teaching-learning process.

The authors wish to acknowledge and express appreciation to Mrs. Kitty Farnell, Science Coordinator of School District Five of Lexington and Richland Counties; Dr. Debra Hamm, Director, Elementary and Secondary Education, Richland School District Two; and Dr. Danny McKenzie, Associate Professor of Elementary Education, University of South Carolina, for their assistance in identifying resources used to prepare this report.

Comments or suggestions concerning the information of this report are welcome and may be directed to the authors at the Center.

## Introduction

The basic skills needed to function in an informational-technological society are changing as the 21st century nears. Skills in reading, writing, and arithmetic are no longer sufficient. New basics such as communication skills, problem-solving skills, and scientific and technological literacy are needed by all students in order to successfully work and live in our changing world (American Association for the Advancement of Science, 1989; National Science Board Commission on PreCollege Education in Mathematics, Science, and Technology, 1983). The field of science education is undergoing changes to bring about reforms in science, mathematics, and technology that respond to current and emerging needs. Recent advances in research on cognition (Resnick, 1987) and reform efforts in science curriculum and instruction (AAAS 1989; Brunkhorst, 1991) are influencing the way in which science is taught and assessed in the classroom.

## Effective Science Learning

Science education encompasses three major categories of learning: knowledge, skills, and dispositions (Raisen et al, 1989). The knowledge category includes several broad areas such as factual and conceptual knowledge as well as knowledge about the nature of science and the interrelationships between science, technology and society. The skills category includes various types of skills such as the basic and integrated process skills which includes observing, measuring, classifying, inferring, predicting, and communicating; science laboratory skills which includes manipulating equipment; and problem solving and reasoning skills. The third category of science learning, dispositions, involves the ability and the inclination to apply science knowledge and skills in daily life (Raisen et al, 1989).

How students learn science is not entirely clear but recent research in cognitive processes supports the constructivist's view of learning. This approach sees learners as actively constructing their own knowledge by engaging in learning experiences and integrating new information with prior knowledge (Resnick, 1987). The constructivist approach has gained favor among science educators who believe that students must be given opportunities to acquire knowledge and develop understandings by direct experience with science concepts and theories. Similarly, students must be allowed time to inquire, explore, and investigate scientific concepts and phenomena in order to deepen their understanding. In general, current thinking about effective science learning leads to the conclusion that the science curriculum must change to focus in depth on a few areas of inquiry rather than providing a diffused superficial coverage of a large domain of topics (AAAS, 1989; Raisen et al, 1989).

A very important aspect of effective science learning is the relationship between learning basic facts and concepts and developing problem solving and reasoning skills. According to Raisen et al (1989) learning in science is not hierarchical. These researchers state that "the acquisition of facts and structuring of a knowledge base goes hand in hand with learning how to apply knowledge, how to reason and solve problems" (p. 52). Thus, a student does not have to master a list of facts and terms before being able to investigate or experiment with related concepts and theories. This is an important departure from the traditional hierarchial view of science learning which emphasizes mastery of facts and terms as prerequisite to understanding concepts and theories.

## Recent Reform Efforts in Science Education

Recent reform efforts in science education emphasize two important goals:

1) the development of "scientific literacy" for all students, not just an elite few, and
2) more in-depth coverage of science topics and concepts, the idea that less is more (AAAS, 1989; Brunkhorst, 1989)

Other major reform recommendations are that the science disciplines should be coordinated to foster connections between the sciences and the real world; science should be taught through a hands-on approach allowing students to develop and test hypothesis, observe phenomena, and manipulate laboratory equipment; and that science should be taught through an inquiry approach which encourages a sense of curiosity and investigation (AAAS, 1989; Brunkhorst, 1991; National Association of Educational Progress, 1992).

There are two major reform efforts in science education for the 1990's that incorporate these goals and recommendations for school science. The American Association for the Advancement of Science (AAAS) Project 2061 (1989) and the National Science Teachers Association (NSTA) Scope, Sequence, and Coordination Project (Brunkhorst, 1991) represent long term endeavors to guide and transform school science for all students (K-12).

Project 2061 is designed as a three phase plan for reforming science, mathematics, and technology education. Phase I is described in Science for All Americans (AAAS, 1989) and establishes a conceptual base for reform by spelling out the knowledge, skills, and attitudes all students should acquire from K-12 school experience. Phase II involves the development of alternative curriculum models for use by school districts and states to promote science literacy. Phase III is a long term collaborative effort to implement reforms and recommendations in our nation's schools. (For more information on Project 2061 the reader is directed to Update - Project 2061, AAAS, 1992.)

The NSTA Scope, Sequence, and Coordination Project of Secondary School Science (Brunkhorst, 1991; Haney, 1990) proposes to abandon the traditional sequencing of science courses over grades 7-12 and implement the teaching of biology, chemistry, physics, and earth science each year in a developmentally appropriate fashion. The Scope, Sequence, & Coordination Project allows for two approaches in teaching science. The sciences may be taught as separate subjects with students spending time on each subject each week, or the various fields of science may be integrated into each year of schooling. These approaches to school science are based on the belief that "instruction is more effective when students can study a subject over several year's time rather than spending a whole year in a single subject and doing nothing with the others that year" (Haney, 1990, p. 4).

There are no existing national standards in science that enjoy widespread consensus although NSTA's Scope, Sequence, and Coordination Project and AAAS's Project 2061 have established benchmarks for science education. The National Research Council (NCR) is coordinating the effort to develop national standards for science education that will represent what is important for all students to attain in science education. The goal is to publish the Science Education Standards by the Fall of 1994. The national standards will be developed by teachers and other educators, scientists, and the general public. This document will provide standards for three components of science education: curriculum, teaching, and assessment. The curriculum, teaching, and assessment standards are designed to provide criteria upon which appropriate curriculum, sound teaching practice, and valid judgements of student work can be based.

## Assessment of Science Learning

The current curriculum and instructional reforms in science education have lead to discussions and suggestions for reform in the assessment of science learning. Many of the proposals for science reform call for improvement in assessment of science learning outcomes and science programs (AAAS, 1989; Brunkhorst, 1991). Raisen et al (1989) and Shavelson, Baxter, and Pine (1991) cite several characteristics of desired assessment of science learning.

Characteristics of Science Assessment

1) Probe and capture students' knowledge base, understandings, reasoning and problem solving skills.
2) Incorporate hands-on performance tasks in which students demonstrate laboratory skills and thinking skills.
3) Align assessment with curriculum and instructional reform to foster and encourage reform in the classroom and to provide a meaningful context in which to interpret assessment results.

Unfortunately, results of a study conducted by Gong, Lahart, and Courtney (1990) on current science assessments used in the United States (grades 6-7) showed that typical science tests were generally not consistent with current science education reform goals. For example, typical tests are often composed of items representing a sample of topics within several disciplines. This approach is clearly not conducive to encouraging depth of learning in one topic as advocated by such reform efforts as Project 2061. Moreover, Raisen and Kaser (1989) point out that the standardized tests used for large scale and classroom assessment cannot assess a student's ability to conduct science investigations and a student's growth in science knowledge and performance.

*Alternative Forms of Assessment*

Experts in science education are calling for assessment techniques that support curriculum and instruction and promote exemplary science learning (Baron, 1991; Gong et al, 1989; Raisen, 1989). Pencil and paper instruments such as standardized multiple choice tests most readily assess factual knowledge (Raisen, 1989) but this is only one dimension of science learning. Reliance on paper and pencil methods is not consistent with efforts to teach science through inquiry or to test student knowledge of science process skills. Current research indicates that alternative forms of assessment such as performance-based assessment, portfolios, and concept maps are being used in classrooms to assess students' knowledge, skills, and dispositions in science. Each of these forms of assessment will be examined.

*Performance Assessment*

Performance-based assessments vary in definition, structure, and administration but all require the student to actively "do" science. For example, one activity may require a five minute experience measuring water temperature while another may require a week long investigation in high school chemistry. In both cases the teacher can assess what the student actually knows and can do by observing an actual performance or behavior, by collecting information through interviews or conversations with the student, and by looking at products of student work. Baron (1991) describes three characteristics of science performance assessments.

<u>Characteristics of Performance-based Assessment in Science</u>

- The task requires active student involvement to solve.
- Students produce the solutions rather than recognize them from a list of choices.
- Students construct and communicate responses rather than recall from a text.

Performance assessments in science can be based on performance tests or tasks, open-ended questions, student journals, and computer simulations. Each method will be described below. A discussion of advantages and disadvantages of using performance-based assessment in science classrooms will follow the explanations of the various types of assessment.

<u>Performance tests/tasks.</u> Performance tests are hands-on activities in the science classroom in which students actually do a task related to the science instruction. Performance tasks are common laboratory settings where students move from station to station manipulating instruments and performing tasks related to specific processes. For example, Shavelson, Baxter, and Pine (1991) describe a hands-on assessment with fifth and sixth graders in which students conduct an investigation to determine which one of three brands of paper towels soaks up the most/least water. The California Assessment Program (1990) field tested a performance task in which students were asked to complete an electrical circuit using items such as plastic spoons, washers, pennies, and rubber bands. Performance tasks

such as these may be scored by simply using a checklist and "checking off" specific actions or a more elaborate system may be developed. A scoring procedure was developed to assess the performance on and results of the paper towel experiment. The problem statement and scoring form for the hands-on paper towel investigation are contained in Appendix A.

Baron (1991) describes an approach used in Connecticut's performance assessment tasks as combining individual work at the beginning and end of the task with group work during the middle. For example, at the beginning of the task, each student's prior knowledge of the relevant scientific concept/procedure is determined by asking the student to give an initial impression, estimate a solution, design a preliminary study and/or list questions about the concepts being assessed. In the second phase, students work in teams to produce a group product for the specified task. During this time a variety of assessment techniques such as journals, checklists, portfolios, student logs, and oral interviews are used to provide evidence that students are learning and understanding. At the end, a related task that assesses the same content and processes as the group task is administered to students individually to determine what students learned from the group experience.

Open-ended questions. Open-ended questions allow students to communicate their understanding of a topic by writing out their response in their own words. This method permits students to justify their reasons for choosing an answer to a question or explain their logic or thought processes. Thinking and reasoning skills are emphasized and thus become valued by students when open-ended questions are used in science classrooms. The following activity from the Massachusetts Educational Assessment Program (1989) illustrates an open-ended question in science assessment.

> "A person wants to determine which of two spot removers is more effective. Describe in detail an experiment the person might perform in order to find out which spot remover is better for removing stains from fabrics." (Badger & Thomas, 1989, p. 13).

Student journals. Student journals and other reflective writing techniques have become successful means of self-assessment of scientific beliefs, values, and understandings (Dana, Lorsback, Hook, & Buiscoe, 1991). Students can be given an enabler such as a specific question or a sentence stem in order to stimulate and focus their thinking. Students may review their journal entries to asses their growth of knowledge, understanding, and attitudes in science. Journals may also be used as a means of dialogue between student and teacher about scientific issues and concerns.

Computer simulations. Advances in computer technology have made possible the use of computers for monitoring students' knowledge and abilities in science. The use of computer simulations and telecommunications have opened up many opportunities for instruction and assessment in science classrooms (Raisen et al, 1990).

Computer simulations may be classified as two types, passive or active. Students observe scientific phenomena in passive simulations, and by seeing and hearing the simulation they gain a richer understanding of what actually occurs and why. As stimuli for assessment, passive simulations can be used to ask students to explain and/or predict related phenomena. Active simulations provide a hands-on learning experience in which students can alter variables and observe the resulting effects. Students have the opportunity to solve complex problems set in real-world contexts. As an assessment procedure, teachers may monitor student understanding by programming the computer to record the strategies and information that the students use in trying to solve the problem. In a study conducted by Shavelson et al (1991), two computer simulations were developed to replicate actual performance assessments for an electronic circuit activity and an investigation using sow bugs. The results of their study with over 300 fifth and sixth grade students suggested limitations in the exchangability of the simulations for the hands-on performance assessments.

The use of electronic telecommunications allows students in different locations to work together on a science problem or investigations. Raisen et al (1990) reports an activity where students in middle schools in various countries and at different altitudes record and report the temperatures at which water boils and share data with each other via electronic communications.

*Benefits and Limitations of Performance-based Assessment*

Performance-based assessment approaches offer many advantages in measuring different aspects of science achievement. Unlike pencil-and-paper multiple choice tests, performance assessments permit valid assessments of affective dimensions such as attitudes and dispositions about science. In addition, performance tasks allow teachers to observe and assess students as they perform science process skills such as observing, classifying, measuring, inferring, predicting, experimenting, and formulating models. Performance assessment also provides opportunities for all students to perform by allowing a variety of approaches and responses to the tasks. The National Association on Educational Progress (NAEP) (1987) conducted a study involving eleven hands-on activities designed for instructional and assessment use in science and mathematics. These tasks ask students to solve problems, conduct investigations, and respond to questions using materials and equipment. Various tasks were administered as group activities requiring open-ended paper/pencil responses to questions; station activities required students to manipulate equipment and investigate relationships; and complete experiments were administered to individual students. Findings from the pilot data indicted that students responded well to the tasks and that teachers were supportive and willing to incorporate more hands-on instruction and assessment in their classrooms. Most importantly, this NAEP study indicated that "conducting hands-on assessment is feasible and extremely worthwhile" (p. 9). Bullock, Collins, Marshall, and Steiner (1991) recommend performance based assessment as advantageous at the district level for evaluating students' process skills and to encourage alignment of curriculum and instruction. These authors state that training and materials can be more efficiently provided at the district level. They point out that a sufficient pool of performance based tasks must be available to prevent narrowing of the curriculum and that provisions must be made to equate performance on different activities from year to year.

Although performance assessments are practical and useful for measuring aspects of science achievement not easily measured by multiple choice tests, there are limitations that should be addressed. Shavelson et al (1991) found in their study of three hands-on investigations with fifth and sixth graders that some students perform well on one task and others perform well on another task. Thus the inference one would make about a student could depend on which specific task was used. These researchers conclude that a substantial number of assessment tasks must be used in order to validly generalize from students' performances to the science domain of interest.

Baron (1991) points out that during a performance assessment the extent of new construction of knowledge and new thinking that students must do depends upon the nature of the task and the prior curriculum and instructional experiences of the student. Shavelson et al (1991) agree and also state that some students may be disadvantaged due to their past experience.

In administering performance tests, issues of reliability are of primary concern when the tests are being used for large scale assessment. For example, in order to ensure reliable results, all test sites must be equipped and arranged with a standard set of materials and raters must be trained to score the responses consistently and must have time for scoring. Establishing reliability for performance tasks is more difficult than establishing reliability for pencil and paper tests. Standardizing types of tasks, length of tasks, scoring scales and procedures, and materials are difficult and can comprise reliability.

Cost is another concern of performance assessment. Blank and Selden (1989) list three types of cost associated with performance assessment. These types are costs in designing the hands-on activities, purchasing equipment for the tasks, and administering and scoring the tasks. The cost of implementing performance-based tests is higher than traditional pencil and paper methods, but the increase in expense varies with the complexity of the tasks involved. Several states have asserted that even though additional funds are needed for equipment and administration of the tasks and additional teacher training is needed to implement alternative assessment programs, the impact on curriculum and instruction is worth it (Aschbacher, 1991).

A discussion of several states that are fully implementing or pilot testing performance assessments in science eduction in place of traditional standardized tests is contained at the end of this paper.

Portfolios. Portfolios are a method of authentic assessment conducive to capturing the complexity of science learning. A science portfolio is a collection of student work that serves as documentation of the students' learning in the science classroom over an extended period of time. Students and teachers may review the contents of the portfolio to assess the student's progress toward attaining the learning outcomes.

The purpose of the portfolio is the first issue to consider when using the portfolio in science assessment. The purpose of the portfolio is shaped by both the curriculum and instruction being offered. The process involves establishing the scientific content that is to be learned, the processes and activities that students should be able to do, and the attitudes about science that students should cultivate. Determining these goals and objectives will then shape the types of evidence appropriate for inclusion in the portfolio.

The uses of the portfolios will help determine whether the portfolio should contain all of a student's science work or only selected "best pieces." Collins (1991) suggests that placing all work in the portfolio is reasonable if the portfolio is to be used by students to review and reflect on their growth and progress over time. However, only "best work" samples are more appropriate if the portfolio is to be used to evaluate students and communicate to parents, administrators and the community what the students know and are capable of doing.

Determining the actual design of the portfolio leads to very important questions such as how much evidence should be contained in the portfolio, who should decide the samples to include, and how the contents should be organized. For example, a teacher may establish requirements for the contents of the portfolio such as five laboratory experiments; four journal entries; a response to a computer simulation; and a project incorporating mathematics, technology, and science.

Collins (1991) expresses the value of portfolio assessment in science.

> "This mode of assessment conjures up a different perception of a classroom- a portfolio cultural center where knowledge is constructed communally, individual strengths are developed, and learning becomes an adventure where success is displayed. - Is there any other way to assess science?" (p.309)

Concept Maps. Dana, Lorsback, Hook, and Buiscoe (1991) have adapted concept mapping (Novak & Gowin, 1984) as a method of alternative assessment in science classrooms. Students identify the concepts of a particular topic, connect them in meaningful ways, and thus demonstrate their understandings of the science concepts and processes. For example, in an eighth grade physical science class, students were asked to generate a list of ideas on light and then organize and connect the ideas in a concept map. Through the use of pre- and postinstruction concept maps, teacher may assess student growth over a particular lesson, unit or course in science.

Lomask, Baron, Greig, and Harrison (1992) have developed a method called ConnMap which uses concept mapping to evaluate students' knowledge of science. In this method, students write essay-type responses to open-ended questions. These responses are converted into concept maps by trained teachers and then are compared to ideal concept maps created by experienced science teachers. The student's map is evaluated by determining the number of concepts included in the map (size dimension) and the number of valid connections made between concepts (strength dimension) that match the ideal map. An overall score reflecting the student's understanding is calculated based on a combination of size and strength dimensions. Although the results are preliminary, the ConnMap method appears to be a valid, reliable and feasible method of assessing student understanding of science concepts (Lomask et al, 1992). An example of a student essay response and the related concept map is contained in Appendix B.

### *Current State Activity With Alternative Forms of Science Assessment*

Davis and Armstrong (1991) report that many states are addressing the limitations of single-answer multiple choice tests and are moving toward alternative approaches such as including performance assessments for state science tests. In 1990, California, New York, Illinois, and Massachusetts used performance -based assessments in their state tests and Connecticut is moving toward exclusive reliance on performance measures in 1992 (Davis and Armstrong, 1992). New York is assessing students' attitudes toward science and Connecticut is also looking at similar plans.

#### *California*

In 1990 California developed a new state science framework which emphasized a thematic approach to science. This framework focuses on aligning assessment with instruction in both content and format. New science tests have been developed in both the sixth and twelfth grades which include multiple choice items, open-ended questions, and performance tasks. The open-ended questions for the sixth grade require students to create hypotheses, design scientific investigations, and write about social and ethical issues in science. The format for the performance assessment consists of five stations which take approximately ten minutes each to complete. The 1991 test included 35 tasks which were randomly sampled and assigned in sets of five to students (Davis and Armstrong, 1991). The performance tasks focus on the science process skills used in life, physical, and earth sciences. In the spring of 1990, California piloted a new testing program called the California Golden State Exams. In biology and chemistry the examinations include conceptually based multiple choice questions, open-ended questions, and performance tasks in a laboratory setting. The examinations are optional and students receive special endorsement on their diplomas.

#### *New York*

In 1989 and 1990, New York administered performance tests in grade four. Students were asked to measure, compare, create, predict, and infer. A multiple choice test addressing factual knowledge and concepts and skills of inquiry and data interpretation was also administered. In addition, districts could elect to survey students on attitudes toward science.

#### *Connecticut*

In 1989 the Connecticut State Department of Education received a grant from the National Science Foundation to form the Connecticut Multi-State Performance Assessment Collaborative Team (CoMPACT). This organization established criteria for effective tasks and developed performance tasks to be used in 1989-1990 science classes. More recently, Connecticut used the criteria and the objectives in science from the Connecticut Common Core of Learning Program to emphasize problem solving and investigations in science assessment in grades 7-12. The assessments consist of group tasks that may last up to one week each. Baron states, "Students will be scored on the content of their solutions, the processes used to arrive at them, the interpersonal and communicative skills demonstrated by the group, and the manifestation of attitudes such as intellectual curiosity." (as cited in Davis and Armstrong, 1991, p. 142).

*Massachusetts*

Beginning in the 1992 assessment, Massachusetts administered open-ended questions to all students in science in the fourth, eighth, and twelfth grades. Questions dealing with two distinct aspects of science comprehension - science inquiry and scientific concepts - were asked. An attempt was made to place all questions in a real-world context. The Massachusetts Department of Education has published a series of reports entitled "On Their Own" which describe and discuss results of the assessments using open-ended questions.

## References


American Association for the Advancement of Science. (1989). Science for all americans. A project 2061 report. Washington, DC: AAAS.

American Association for the Advancement of Science. (1992). Update - project 2061. Washington, DC: AAAS.

Aschbacher, P.R. (1991). Performance assessment: State activity, interest, and concerns. Applied Measurement in Education, 4(4), 275-288.

Badger, E., & Thomas, B. (1989). On their own: Student responses to open-ended tests in science. Quincy, MA: Massachusetts Department of Education.

Baron, J.B. (1991). Performance assessment: Blurring the edges of assessment, curriculum, and instruction. In G. Kulm and S. Malcom, (Eds.), Science assessment in the service of reform (pp. 247-266). Waldorf, MD: AAAS.

Blank, R. & Selden, R. (1989, January). Alternative methods for assessing science: Report to the states. Conference report of the state assessment directors and state science supervisors. Washington DC: National Science Foundation.

Brunkhorst, B. (1991). Every science, every year. Educational Leadership, 49(2), 36-38.

Bullock, J.T., Collins, C.B., Marshal, J.E. & Steiner, H.E. (1991, March). Measuring science process skills: Where we have been and what's in the future. A paper presented at the National Science Teachers annual convention, Houston, TX.

California Assessment Program. (1990). New directions in CAP science assessment. Sacramento, CA: California Department of Education.

Collins, A. (1991). Portfolios for assessing student learning in science: A new name for a familiar idea? In G. Kulm and S. Malcom, (Eds.), Science assessment in the service of reform (pp. 291-300). Waldorf, MD: AAAS.

Dana, R.M., Lorsback, A.W., Hook, K., & Buiscoe, C. (1991). Students showing what they know: A look at alternative assessments. In G. Kulm and S. Malcom, (Eds.), Science assessment in the service of reform (pp. 331-337). Waldorf, MD: AAAS.

Davis, A., & Armstrong, J. (1991). State initiatives in assessing science education. In G. Kulm and S. Malcom, (Eds.), Science assessment in the service of reform (pp.127-146). Waldorf, MD: AAAS.

Gong, B., Lahart, C. & Courtney, R. (1990, April). Current state science assessments: Is something better than nothing? Paper presented at the meeting of the American Educational Research Association, Boston, MA.

Haney, R.E. (1990, October). Back to the future of science and mathematics education. Paper presented at the meeting of the School Science and Mathematics Association, Cincinnati, OH.

Kulm, G. & Malcom, S. (Eds.). (1991). Science assessment in the service of reform. Waldorf, MD: AAAS.

Lomask, M., Baron, J.B., Greig, J., & Harrison, C. (1992, March). ConnMap: Connecticut's use of concept mapping to assess the structure of students' knowledge of science. A symposium presented at the meeting of the National Association of Research in Science Teaching, Cambridge, MA.

National Association of Educational Progress. (1992). Science framework for 1994: NAEP science consensus project. Washington, DC: National Assessment Governing Board.

National Association of Educational Progress. (1987). NAEP: Learning by doing: A manual for teaching and assessing higher order thinking in science and mathematics. Princeton, NJ: Educational Testing Service.

National Science Board Commission on PreCollege Education in Mathematics, Science, and Technology. (1983). Educating America for the 21st century. Washington, DC: National Science Foundation.

National Research Council. (1992, October). National science education standards project update. Washington, DC: NRC.

New York State Department of Education. (1988). New York state elementary science program evaluation test: Information on test administration. Albany: State University of New York.

Novak, J.D. & Gowin, D.B. (1984). Learning how to learn. New York, NY: Cambridge University Press.

Raizen, S., Baron, J.B., Champagne, A.B., Haertel, E., Mullis, I.V.S., & Oakes, J. (1990). Assessment in science education: The middle years. Andover, MA: The National Center for the Improvement of Science Education, The Network.

Raizen, S., Baron, J.B., Champagne, A.B., Haertel, E., Mullis, I.V.S., & Oakes, J. (1989). Assessment in elementary science education. Andover, MA: The National Center for the Improvement of Science Education, The Network.

Raizen, S.A. & Kaser, J.S. (1989). Assessing science learning in elementary school: Why, what, and how? Phi Delta Kappan, 70(9), 718-722.

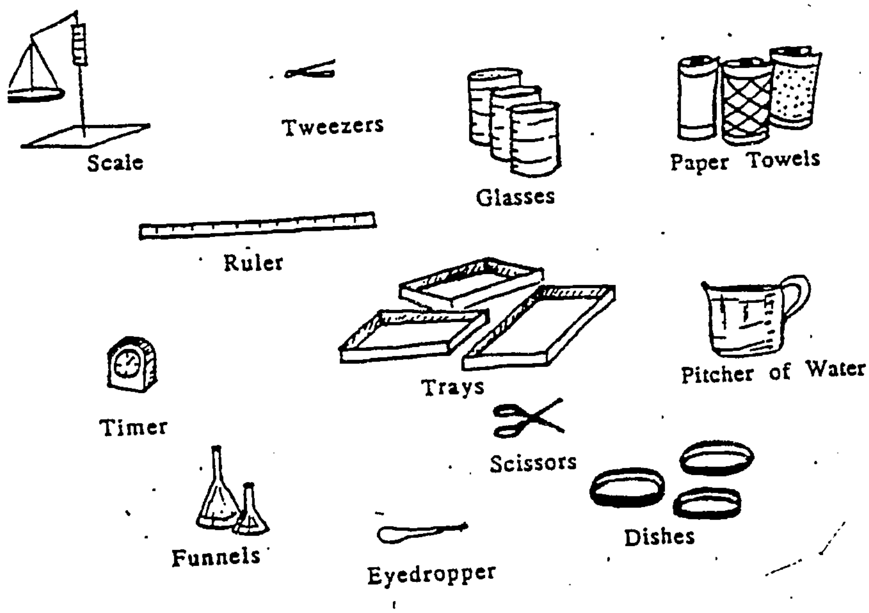
Resnick, L.B. (1987). Education and learning to think. Washington, DC: National Academy Press.

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

## Appendix A

## Hands-On Paper Towels Investigation

Problem Statement: You have three different kinds of paper towels in front of you and some equipment for doing scientific experiments.




Find out which paper towel can hold, soak up or absorb the most water.

Find out which paper towel can hold, soak up or absorb the least water.

Source: Shavelson et al, 1991, p.352.

Appendix A (continued)

## Hands-On Paper Towels Score Form

Student ____________________ Observer ________________________ Score ____

**1. Method for Getting Towel Wet**

| A. Container | B. Drops | C. Tray | D. No Method |
|---|---|---|---|
| Pour water in/put towel in<br>Put towel in/pour water in<br>1 pitcher or 3 breakers/glasses | | Towel on tray/ pour water on<br>Pour water on tray/put towel in | |

**2. Saturation** A. Yes B. No C. Controlled (Same amount of water - all towels)

**3. Determine Result**
- A. Weigh towels
- B. Squeeze towel/measure water (weight or volume)
- C. Measure water in/out
- D. Count # drops until saturated
- E. Irrelevant measurement (ie. time to soak up water, see how far drops spread out, feel thickness)
- F. Other ____________________________

**4. Care in saturation and/or measuring.** Yes No A little sloppy (+/-)

**5. Correct result** Most Least

| Grade | Method | Saturate | Determine Result | Care in Measuring | Correct Answer |
|---|---|---|---|---|---|
| A | Yes | Yes | Yes | Yes | Both |
| B | Yes | Yes | Yes | No | One or B |
| C | Yes | Controlled | Yes | Yes/No | One or B |
| D | Yes | No or | Inconsistent | Yes/No | One or B |
| F | Inconsistent or | No and | Irrelevant | Yes/No | One or B |

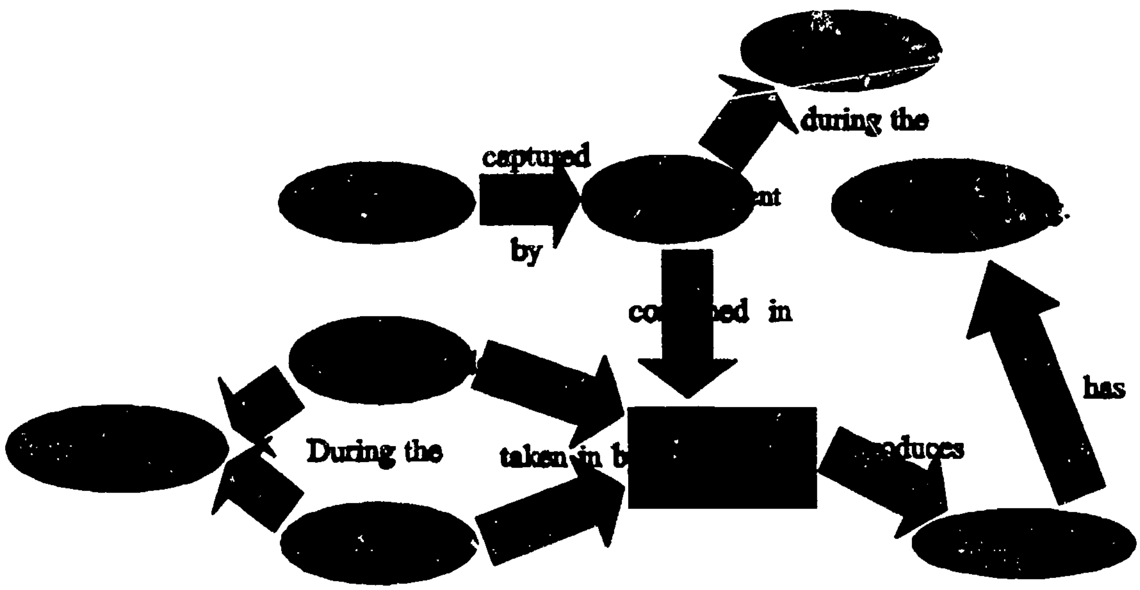
Source: Shavelson et al, 1991, p. 353.

## Appendix B

Question: Describe the types of energies and materials that are involved in the process of a growing plant and explain how they are related.

Student Response: "The sun's energy is taken in by green chloroplasts in the light phase and is mixed with water and energy from the sun to produce glucose, which is the energy stored in the plant."

**Concept Map:**



## ADDITIONAL PUBLICATIONS BY CEASL

ASSESSMENT OF STUDENT LEARNING IN EARLY CHILDHOOD EDUCATION

ASSESSMENT OF STUDENT LEARNING IN FINE ARTS

ASSESSMENT OF STUDENT LEARNING IN FOREIGN LANGUAGE

ASSESSMENT OF STUDENT LEARNING IN LANGUAGE ARTS

ASSESSMENT OF STUDENT LEARNING IN MATHEMATICS

ASSESSMENT OF STUDENT LEARNING IN SOCIAL STUDIES

RESOURCE DIRECTORY OF ASSESSMENT MATERIALS

To obtain copies of the documents listed above please write:

South Carolina Center for Excellence
in the Assessment of Student Learning (CEASL)
College of Education
University of South Carolina
Columbia, SC 29208