DOCUMENT RESUME

ED 415 271 TM 027 978

AUTHOR Enger, Sandra K.
TITLE Linking Instruction and Assessment in Science: Science Learning Opportunities and Student Performance on a Set of Open-Ended Science Questions.
PUB DATE 1997-11-00
NOTE 21p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Memphis, TN, November 12-14, 1997).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Critical Thinking; *Educational Assessment; Grade 9; Graphs; Inquiry; *Instructional Effectiveness; Intermediate Grades; Junior High Schools; Middle Schools; Rural Schools; Science Education; *Science Instruction; Science Teachers; Sciences; Scientific Attitudes; Scoring; Surveys; Tables (Data); *Transfer of Training
IDENTIFIERS *Middle School Students; Open Ended Questions

ABSTRACT

The link between instruction in middle school science and assessment in ninth grade science in a small rural midwestern school district was investigated. Science students, in middle school grades 6-8 (917) and their 13 science teachers responded to surveys used to characterize science learning opportunities in the middle school classes. Ninth grade students (309) responded to a set of open-ended science questions that were developed from a standardized science test. The survey data provided a contextual framework for interpretation of ninth grade student performance on the open-ended science questions. The open-ended questions included questions for which students were required to graph and interpret data, write conclusions, identify control variables, and judge the validity of information. These open-ended questions were scored by three raters according to a scoring rubric. Middle school students did report having science learning opportunities to practice science as inquiry, and their teachers reported that much of the science inquiry practice was contextual. At least 90% of the ninth graders were from the same school district, but when practice in controlling data and making predictions was reported, student data from the open-ended questions suggested minimal transfer of these skills. The interpretation and use of graphic information and the actual graphing of data were areas that were problematic for ninth graders, as was the ability to provide reasons in support of answers. Recommendations based on this study include using a variety of contextual settings for framing science inquiry practice and more attention on discussing, reading, and writing in the content area. (Contains 8 tables and 12 references.) (Author/SLD)

Linking Instruction and Assessment in Science: Science Learning Opportunities and Student Performance on a Set of Open-Ended Science Questions

Sandra K. Enger
The University of Alabama in Huntsville
Institute for Science Education
Morton Hall 122
Huntsville, AL 35899

engers@email.uah.edu
205-890-6670

Paper presented at the annual meeting of the
Mid-South Educational Research Association (MSERA)
November 12-14, 1997
Memphis, TN

Linking Instruction and Assessment in Science: Science Learning Opportunities and Student Performance on a Set of Open-Ended Science Questions

Abstract

The link between instruction in middle school science and assessment in ninth grade science in a small, rural, midwestern school district was investigated. Nine hundred seventeen middle school (grades 6-8) science students and their thirteen science teachers responded to surveys used to characterize science learning opportunities in the middle school science classes. Three hundred nine ninth grade students responded to a set of open-ended science questions which were developed from a standardized science test. The survey data provided a contextual framework for interpretation of ninth grade student performance on the open-ended science questions. The open-ended questions included questions for which students were required to graph and interpret data, write conclusions, identify control variables, and judge the validity of information. These open-ended questions were scored by three raters according to a scoring rubric.

Middle school science students did report having science learning opportunities to practice science as inquiry, and their teachers reported that much of the science inquiry practice was contextual. At least 90% of the ninth grade students responding to the open-ended questions had been middle school students in the district and would presumably have had similar science experiences in their middle school years. While practice was reported in controlling variables and making predictions, student data from the open-ended science questions suggest minimal transfer of these skills. The interpretation and use of graphic information and actual graphing of data were areas that were problematic for ninth grade students. The ability to provide reasons in support of answers was also problematic across questions. Major recommendations based on both survey and open-ended question data included using a variety of contextual settings for framing science inquiry practice and more attention focused upon discussing, reading, and writing in the content area.

Introduction

Assessment should play an integral role in guiding classroom science instruction, wherein instruction and assessment interact in a dynamic state. The purposes of both instruction and assessment should guide this interplay which is targeted toward assessing student learning outcomes. If the science assessment being conducted is designed to assess the students' educational development, not only would the performance on the assessment be of interest, but some attempt to examine the learning opportunities that students bring to the assessment forum would be warranted.

What students experience in the classroom is largely determined by the teachers' instructional goals and objectives; the knowledge and processes teachers make available; the books, materials, and equipment teachers use; the classroom activities teachers arrange; the quality of the teachers' background, training, and experience; and the support and resources available to teachers (Oakes, 1990). Tests can be influential in deciding what content and skills to teach and control the opportunity to learn the full curriculum (LeMahieu & Leinhardt, 1985). While this control of the opportunity to learn the full curriculum may have a negative tone, Resnick and Resnick (1992) note that assessments can serve as positive tools in creating schools that are capable of teaching students to think.

Resnick and Resnick also view assessment as a means to reform education, and Resnick (1994) sees assessments as setting standards to which students and teachers can direct their efforts. Implicit in these views are the understandings that the assessments must be psychometrically sound and in alignment with the purposes for which the assessments are intended.

A case can be made for the inclusion of standardized measures in the assessment of students. Selection of appropriate standardized measures which have normative data can provide information about relative strengths and weaknesses of school programs, and these tests allow for comparisons with a broad and representative student population (Feldt, Ansley, Forsyth, & Alnot, 1994). Nitko (1993) takes a position that educational tests, whether internally or externally developed, can be potent educational tools that can enhance the instructional process. For this to occur, information about the science learning experiences and the assessment information must be linked, and an instruction-assessment feedback loop must exist.

Study Purpose

A major purpose of this study was to examine ninth grade student performance on a set of open-ended science questions developed from a standardized test, the Iowa Tests of Educational Development (ITED). Also of interest in the study were the kinds of science inquiry learning opportunities fostered by middle school science teachers and experienced by middle school students in grades six through eight. This information relative to the science inquiry learning opportunities provided a context in which to examine ninth grade student performance on the set of open-ended science questions. The background information also provided a means for comparison of student and teacher perceptions and could target areas for future instructional attention.

Study Samples

The study samples were drawn from a small, midwestern school district which had a seven year affiliation with a National Science Foundation (NSF) funded, science education reform project. Nine hundred seventeen middle school students in grades six through eight responded to a questionnaire about their science inquiry learning opportunities. Thirteen middle school science teachers were also surveyed with respect to the kinds of science inquiry learning opportunities that were supported in their classrooms. Three hundred nine ninth grade students completed a set of open-ended science questions developed from the ITED. At least 90 percent of these ninth grade students had completed three years of science in middle school classrooms in this district where the teachers had been affiliated with the project. This science education reform project was guided by a constructivist philosophy, and the science teachers in these classrooms worked to facilitate learning environments informed by constructivism.

Instrumentation

Student and Teacher Science Inquiry Questionnaires

Separate student and teacher questionnaires were developed to ascertain perceptions about the science inquiry learning opportunities that existed in the middle school science classrooms. Since the background knowledge and reasoning skills that students bring to a test strongly influence the ability to understand the material on the test (Feldt et al., 1994), the reported experiences provided a context in which to view ninth grade student performance. The questionnaires were developed to collect information about the kinds of classroom experiences, assignments, and material/equipment usage relative to science inquiry as described by the National Science Education Standards (NSES) (National Research Council [NRC], 1996) and in Benchmarks for Science Literacy (AAAS, 1993). Both the teacher and student forms of the questionnaires were reviewed by the district curriculum director and the twenty-one district science teachers, and suggested revisions from this expert panel were incorporated in the final instruments.

Iowa Tests of Educational Development (ITED)

In addition to teacher-designed classroom assessments, students may encounter other performance measures such as the ITED. A primary reason for using a standardized achievement battery like the ITED is to use the resulting information to improve instruction (Feldt et al., 1994). The student results from the ITED provide one perspective on student achievement that may be used in concert with other measures of student learning, and the science section of the ITED can be used to examine student performance, especially with respect to the ability to think critically about diverse kinds of scientific information.

The items from the science section of the ITED are classified into three categories which are interpreting information, analyzing experimental procedures, and analyzing and evaluating information. The kinds of abilities that students are called upon to use in the ITED align with science as inquiry as delineated in the NSES. Experiences to facilitate and develop understandings of science inquiry are to be a component of science learning from kindergarten through grade twelve (NRC, 1996). As set out in the NSES, understandings about scientific inquiry should include scientific investigations in which students are involved in asking and answering questions and comparing their answers with what scientists know about the world.

The expectation exists that opportunities that foster students' abilities in their development of descriptions, explanations, predictions, and models using evidence will be facilitated. These kinds of opportunities should engage students to think critically and logically to help build the relationships between evidence and explanations. Also, mathematics is to be incorporated in all aspects of scientific inquiry, and communication of procedures and explanations is expected (NRC, 1996).

Open-Ended Question Development, Administration, and Scoring

A set of open-ended questions was developed from a subset of multiple-choice items from the science test of the ITED to examine student performance on these items in an open-ended format. The open-ended questions were comprised of four contextual question sets, and each question set included multiple components. After pilot testing and revision, these open-ended questions were administered to 309 ninth grade students in the selected midwestern high school. The open-ended format included questions in which the students were required to graph and interpret data, write conclusions, identify control variables, and judge the validity of information. These open-ended questions were administered over two consecutive days during 50 minute class periods. A scoring rubric for the open-ended response format was developed and used by three trained raters to score the ninth grade student responses.

Results and Discussion

The focus of the results and discussion in this section are based primarily on student responses to two of the open-ended question sets. A description of the two question sets is provided along with the components that were scored. Descriptive questionnaire information from middle school students and middle school teachers relative to the context of the questions is also presented.

Open-Ended Question Description: Effect of Temperature on Dissolving

Students were asked to respond to an open-ended question set that presented data from an experiment about the relationship of water temperature and sugar dissolution time. A data table provided information about the glass number, five different water temperatures, and sugar dissolving times at each temperature. Students were asked to graph the tabled data and then estimate the times required for one spoonful of sugar to dissolve at each of two different temperatures. One temperature was within the tabled temperature ranges reported, and the second was outside the data range. A question was then posed about confidence in the time estimation for sugar dissolving at these two temperatures, and students were asked to explain why they were confident about their selected estimation.

Teacher and Student Survey Responses: Experience with Data Representation

The experiences that the ninth grade students brought with them to the assessment would likely contribute to their performance. To gain insight into the ninth grade science learning experiences, both middle school students and middle school teachers responded to surveys that asked about science learning opportunities in their classrooms. At least 90% of the ninth grade students had been middle school students in the district and would likely have had similar science experiences. Middle school teachers were asked to make judgments about student practice relative to the kinds of skills and abilities that could relate to graphing.

Middle school students were asked to respond to survey items with respect to their judgment of the frequency with which they had experience with skills and abilities that are relative to graphing, and these student responses are presented in Table 1. Approximately, one-fourth of the students reported that they seldom graph data.

(Insert Table 1 about here)

Middle school teacher frequencies for reported practice in activities related to graphing are provided in Table 2. As shown in Table 2, 12 of the 13 middle school teachers reported that students have either on-going practice or practice in certain contexts with data collection. Practice with graphing data varies more across the 13 teachers with practice in certain contexts reported most frequently, and four teachers noted either limited or no practice in graphing data.

(Insert Table 2 about here)

Gender differences are often of interest in student performance and in particular, in areas like science and mathematics. As shown in Table 3, when contrasted by gender, survey information supplied by middle school students in response to the kinds of skills and abilities relative to graphing showed significant differences for perceptions for two of three components related to writing. Females perceived that they more frequently wrote down their own observations from an experiment and wrote about the experiments that were done in a notebook, log, or journal. While these contrasts are statistically significant, they may not be of practical significance based upon the sample sizes, since the contrast groups are large enough that minor differences can become statistically significant.

(Insert Table 3 about here)

Scoring Procedures and Response Comments

A rubric was used by three raters to score this question set on the basis of a nine-point dichotomous scoring scheme. Since students were asked to graph the tabled data, a ruler and a sheet of graph paper on which to construct the graph were provided for each student. The elements included in the scoring guide included: placement of an appropriate title on the graph; placement of the independent variable on the x-axis and the dependent variable on the y-axis; appropriate scaling of both the x and y axes; type of graph used to represent data; ability to make predictions of two data points based upon data presented; student confidence in a datum prediction; and a reason for confidence in this prediction.

Table 4 provides the score point information on the dichotomously scored components and gives the interrater percent agreement for this question by each element assessed. Cronbach's alpha, a measure of internal consistency, was calculated for each of the three raters for all open-ended question sets, and alpha values were .86, .85, and .83. The mean on this nine point item was 5.88 (S.D. 1.66) for females, and for males was 5.77 (S.D. 1.75). The t-value (-.57) was not significant.

(Insert Table 4 about here)

Placement of a title on a graph may have been an oversight by students based upon their background experience with graphing. The placement of the manipulated or independent variable, temperature in this data set, on the x-axis and the placement of the dependent variable, time, on the y-axis would be matter of convention. Whatever the chosen data orientation, students were more successful at scaling the axes.

Because the time and temperature data are both continuous data, a line would be considered a more appropriate representation of the time and temperature relationship than a bar graph. When student papers were examined, some students had plotted only points but did not connect these data points with an actual line.

Students were more successful at making time estimates for sugar dissolution for temperatures not included in the tabled data than for some of the other elements scored. While students may have used their time and temperature graphs to make these estimates, evidence from student work indicated that they also made calculations based upon using the tabled time and temperature data.

When students were queried about the temperature estimate for which they were most confident, 43% of the students correctly identified the temperature (temperature 1) that was within the data range. Thirty-five percent of the students further supported this estimate with answers that indicated that this temperature was included within the temperature range for which tabled data were presented. The word confident was also interpreted by a number of students in a context that they were confident about either or both estimates because they knew how to work the problem or had worked the problem carefully. The idea of having greater confidence in a datum prediction within the range for which data were collected was not overtly stated by students, and a more everyday use of the idea of confidence was presented by students.

The background that middle school students report about the frequency of experience of seldom graphing data from their experiments may influence students' ability to manage graphic data. The assessment information from this open-ended question can be linked to, or least questioned in, the context of the reported instructional experience. When student performance is viewed in the context of survey information, skills related to graphic representation may be an area in which more instruction could be targeted.

Information from student papers on the open-ended question set also brought to attention some other areas in which students may benefit from discussion and practice. These included mechanics such as placing title on a graph, assignment of variables to the axes, and selection of an appropriate graph for data representation. On a multiple-choice item from the ITED, in which students from this same sample were asked to select the graph that matched tabled data, only 22.5% students selected the correct response from four foils. The middle school teachers also reported less practice in the area of graphing.

As noted earlier, it also appears that students did not use the graph to answer other parts of the question but instead relied on manipulating the data from the data table. While no students used the word extrapolation, some students did express this idea. The concept of reasons for having confidence in data may be an area that would be of value to discuss with students and may be an idea that has not been brought to students' attention.

Open-Ended Question Description: Effect of Temperature on Cricket Chirps

In another of the open-ended questions sets, students were asked to read a research summary about the relationship between the rate of cricket chirping and ambient temperature. A graph of the data was included with the contextual paragraph, but students were not actually asked to graph the data. The kinds of skills and abilities that were used in answering the temperature-sugar dissolution question would also appear to be relative to answering the temperature-cricket chirping question. Student reports of practice related to the cognitive demands of the cricket question are presented in Table 5, and teacher reported experiences that could relate to student proficiency in answering the cricket question are presented in Table 6.

(Insert Table 5 about here)

(Insert Table 6 about here)

Scoring Procedures and Response Comments

The nine elements that were scored were: the ability to discern and write the question that researcher was investigating; identification of at least one control variable; determination of chirping rates for each of four different temperatures; selection of a temperature prediction for which the student had the least confidence; a reason for selection of this temperature; and a written statement of the relationship supported by the data for the temperature and rate of cricket chirps. The temperature/cricket chirp question was scored as a nine-point item for which each of nine elements was dichotomously scored. Table 7 gives score point information for the elements assessed in the cricket question is reported.

(Insert Table 7 about here)

The mean for this nine (9) point item was 5.10 (S.D. 2.02), and when contrasted by gender the mean for females was 5.18 (S.D. 2.08), and for males was 5.06 (S.D. 1.99). The t-value (-.50) was not significant. Students had a tendency to set out words which were not in a sentence format, and it became the task of the rater to imply the intended meaning. The research question did not have to be stated in question form, but the student answer should have indicated with some clarity that the researcher was investigating the relationship between temperature and the rate of cricket chirping.

Only about one-third of the students were able to identify that the time of day for data collection was a control variable. A range of chirping rates was accepted for each of the temperatures for which students were to make predictions. Based upon the score-point information, students apparently had more difficulty with temperatures two (2) and three (3). Given the nature of the graph, approximating a line of best fit may have been an approach to make more accurate predictions, with lines then drawn to the axes to help locate the coordinates. Other demands of the problem may have been related to lack of grid lines and the necessity to scale the axes to locate a temperature and a chirping rate.

About one-third of the students gave a reason for confidence that was agreed upon by all three raters. Again, many students did not appear to have the idea of confidence in predictions in the context of scientific work. Students (81.2%) were better able to write a conclusion for the findings of the researcher. Teachers did report that students have practice drawing inferences, in developing science hypotheses, and posing researchable questions. However, practice that is context-specific may not transfer to another contextual setting.

Gender Differences in Performance

Gender differences in student performance are of interest in assessment especially in science and mathematics. Only two question sets have been presented in some detail, but when student performance was contrasted across all open-ended question sets, no significant differences were noted by gender. As shown in Table 8, for this same ninth grade sample, when performance on a multiple-choice form of the ITED was contrasted by gender, a significant difference by gender was noted.

(Insert Table 8 about here)

Becker and Forsyth (1994) reported that a study conducted by Stroud and Lindquist in 1942 found gender differences on the Iowa Every-Pupil Tests of Basic Skills. In the areas of language, work study skills, and reading females maintained a consistent and usually significant advantage over males. Males had a small but insignificant advantage over females in mathematics. The heavier reading load and verbal abilities may favor females on the multiple-choice format of the ITED. Becker and Forsyth also noted that examining male and female performance at different ability levels provided a clearer picture of the relationship between academic achievement and gender than just looking at group averages. Differences in test performance by gender do not always have clear explanations. Burton (1996) found that girls commonly score better on than boys on reading tests, and young women tend to score better than young men when subject matter is not gender-specialized.

Implications

The use of an open-ended question format does provide diagnostic information about student performance. Given a stated purpose of the ITED is the intended use of the test information to improve instruction, the information from the open-ended format added insight into student performance with respect to the strengths in students' answers along with errors that could be used for diagnostic purposes. When Birenbaum and Tatsuoka (1987) compared open-ended versus multiple-choice formats, they found considerable differences between the formats. As they noted, while multiple-choice formats are considerably easier to score, this format may not provide the appropriate information for identifying students' misconceptions. Also, precisely the kinds of errors that students make in relation to the nature of the cognitive demands of the items may be more difficult to assess with the multiple-choice format.

The questionnaire information from students and teachers also helps to interpret student answers on the open-ended question sets. Assessment information from multiple sources can be linked back to inform instruction which can assist students in the development of conceptual understanding in science. In their analyses of the NELS:88 data, Burkam, Lee, & Smerdon (1997) reported that more frequent use of student-focused labs in conducting experiments and writing reports contributed to all students' learning especially in the physical sciences. Based on the information from survey responses from both middle school students and teachers, a recommendation for inclusion of a wider variety of contextual practice in science investigations seems warranted. Attention to graphing skills in the science content area and the analysis of a variety of data representations would be recommendations.

The reading and discussion of a variety of science related materials would be recommended with a focus on discerning the questions being researched and attention given to judging the validity of claims made based upon the reported research. Practice also would be recommended in writing with a focus on explanation and evaluation in the content area. If students are expected to make applications of what they understand and transfer skills and knowledge to other problems, a wider experiential base in science should be considered.

References

American Association for the Advancement of Science. (1993). Benchmarks for science literary. Oxford, NY: Oxford University Press.

Becker, D. F., & Forsyth, R. A. (1994). Gender differences in mathematics problem solving and science: A longitudinal analysis. International Journal of Educational Research, 21(4), 407-416.

Birenbaum, M., & Tatsuoka, K. K. (1987). Open-ended versus multiple-choice response formats—It does make a difference for diagnostic purposes. Applied Psychological Measurement, 11(4), 385-394.

Burkam, D. T., Lee, V. E., & Smerdon, B. A. (1997). Gender and science learning in high school: Subject matter and laboratory experiences. American Educational Research Journal, 34(2), 297-331.

Burton, N. (1996). Have changes in SAT affected women's mathematics performance?. Educational Measurement: Issues and Practice, 15(4), 5-9.

Feldt, L. S., Forsyth, R. A., Ansley, T. N., & Alnot, S. D, (1994). Iowa tests of educational development interpretive guide for teachers and counselors. Chicago, IL: Riverside Publishing.

LeMahieu, P., & Leinhardt, G. (1985). Overlap: Influencing what's taught: A process model of teachers' content selection. Journal of Classroom Interaction, 21(1), 2-11.

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

Nitko, A. J. (1993). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), Educational Measurement (pp. 447-474). Phoenix, AZ: The Oryx Press.

Oakes, J. (1990). Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science. Santa Monica, CA: The RAND Corporation.

Resnick, L. B. (1994). Performance puzzles. American Journal of Education, 102 (4), 511-526.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: new tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), Changing assessments alternative views of aptitude, achievement and instruction (pp. 37-75). Boston: Kluwer.

Table 1. Middle School Students' Report of Experience in Data Collection and Representation

Type of Skill and/or Ability	N	Response as Percent: Very Often	Often	Some-times	Seldom	Never	Mean	S.D.
Setting up a data table when doing activities or experiments	915	34.4	35.2	21.1	6.8	2.5	3.92	1.02
Graphing numbers from their experiments	914	10.7	21.9	32.7	25.6	9.1	3.00	1.13
Writing down their own information from a science experiment	916	19.0	32.5	29.4	14.1	5.0	3.46	1.10
Writing down their own observations from an experiment	917	39.4	34.4	17.5	6.7	1.9	4.03	1.00
Writing about the experiments that are done in a notebook, log, or journal	917	23.4	25.7	23.4	14.6	12.8	3.32	1.32
Making predictions about what will happen before doing activities or experiments	917	35.3	33.5	22.0	7.4	1.7	3.93	1.01

1= Never, 2= Seldom, 3= Sometimes, 4= Often, 5= Very Often

Table 2. Middle School Teachers' Report of Student Practice Related to Graphing

Skill and/or Ability	N	On-going practice	Practice in certain contexts	Limited practice	Don't do this in class
Collecting data	13	7	5	1	0
Graphing data	13	3	6	3	1
Interpreting data from graphs	13	4	5	3	1
Making predictions	13	9	4	0	0
Drawing inferences	13	8	4	1	0
Keeping a laboratory log or journal	13	1	5	5	1

Table 3. Middle School Student Gender Contrasts Relative to Graphing Experience

Type Skill and/or Ability	N_f	Mean Females (SD)	N_m	Mean Males (SD)	t-value
Setting up a data table when doing activities or experiments	446	3.96 (1.00)	466	3.88 (1.05)	1.21
Graphing numbers from their experiments	445	2.96 (1.14)	466	3.04 (1.11)	-1.12
Writing down their own information from a science experiment	445	3.48 (1.07)	467	3.45 (1.13)	.37
Writing down their own observations from an experiment	445	4.13 (.98)	468	3.94 (1.00)	2.93**
Writing about the experiments that are done in a notebook, log, or journal	446	3.45 (1.30)	468	3.21 (1.34)	2.77**
Make predictions about what will happen before doing activities or experiments	446	3.92 (1.01)	468	3.94 (1.01)	-.15

1= Never, 2= Seldom, 3= Sometimes, 4= Often, 5= Very Often

N_f = females, N_m = males

**$p<.01$

Table 4. Effect of Temperature on Dissolving:
Percent Agreement by Three Raters for 1 or 0 Score-Points

Elements Assessed	1 Score-Point Agreement 100%	0 Score-Point Agreement 100%
Title on graph	5.2	88.7
Axes labeled with temperature on x-axis and time on y-axis	55.0	32.5
x-axis scaled	74.8	10.0
y-axis scaled	81.2	11.3
Line graph used	64.4	29.1
Time estimate for temperature 1	81.9	10.7
Time estimate for temperature 2	92.9	4.2
Confidence in time estimate	43.0	47.6
Reason for confidence	35.0	37.9

Table 5. Middle School Student Report of Experience Relative to Temperature and Cricket Chirps

		Response as Percent						
Type of Skill and/or Ability	N	Very Often	Often	Some-times	Seldom	Never	Mean	S.D.
Testing a hypothesis or question in their activities or experiments	908	26.2	34.5	26.2	11.0	2.1	3.72	1.04
Controlling variables when doing lab activities or experiments	913	18.2	30.9	37.3	10.2	3.4	3.50	1.01
Discussing the results from their experiments	914	40.0	33.2	19.9	5.6	1.3	4.05	.97
Reading about the research work that scientists do	914	7.0	12.8	27.5	36.0	16.7	2.57	1.12
Discussing the research work that scientists do	917	4.6	12.4	31.1	34.7	17.2	2.52	1.06
Discussing science articles from newspapers or magazines	916	7.5	13.8	27.7	28.1	22.9	2.55	1.20

1= Never, 2= Seldom, 3= Sometimes, 4= Often, 5= Very Often

Table 6. Middle School Teacher Report of Student Opportunities Relative to Temperature/Cricket Chirp Question

Skill and/or Ability	N	On-going Practice	Practice in certain contexts	Limited Practice	Don't do this in class
Developing science hypotheses	13	4	9	0	0
Posing researchable questions followed with an opportunity to investigate questions	13	2	8	3	0
Controlling variables	13	4	9	0	0
Interpreting data from graphs	13	4	5	3	1
Drawing inferences	13	8	4	1	0
Making predictions	13	9	4	0	0
Reading and discussing the work of scientists	13	1	4	7	1
Using critical thinking	13	9	3	1	0

Table 7. Score Point Information for Temperature/Cricket Chirps

Question III Elements Assessed	1 Score Point Agreement 100%	0 Score Point Agreement 100%
Writing the research question	48.5	28.2
Identification of time variable	31.7	58.6
Chirping rate for temperature 1	80.3	3.9
Chirping rate for temperature 2	33.7	25.6
Chirping rate for temperature 3	22.7	43.7
Chirping rate for temperature 4	50.8	41.7
Confidence in time estimate	52.1	44.0
Reason for confidence	32.7	50.2
Provision of research conclusion	81.2	10.4

N = 309

Table 8. ITED Performance Contrasts by Gender for Ninth Grade Examinees for Multiple-Choice Format

ITED Form K	N_f	Mean (Females)	SD	N_m	Mean (Males)	SD	t -value
Test Score	175	25.58	9.16	159	22.94	9.29	-2.61**

N_f = females, N_m = males
**p<.01

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Linking Instruction and Assessment in Science: Science Learning Opportunities and Student Performance on a set of Open-Ended Science Questions

Author(s): Sandra K. Enger

Corporate Source: University of Alabama in Huntsville, Huntsville, AL 35899

Publication Date: (Nov. 12-14, 1997)

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

[X] Check here **For Level 1 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample sticker shown below will be affixed to all Level 2 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

[ ] Check here **For Level 2 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: Sandra K. Enger

Printed Name/Position/Title: Sandra K. Enger, Assistant Professor

Organization/Address: University of Alabama in Huntsville Institute for Science Education Morton Hall 122, Huntsville, AL 35899

Telephone: 205-890-6670

FAX: 205-890-6922

E-Mail Address: engers@email.uah.edu

Date: 11/13/97

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:

Address:

Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:

Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Assessment and Evaluation
210 O'Boyle Hall
The Catholic University of America
Washington, DC 20064

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2d Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com