DOCUMENT RESUME

ED 387 319 | SE 056 582

AUTHOR Bass, George M., Jr.; Ries, Roger R.
TITLE Scientific Understanding in High Ability High School Students: Concepts and Process Skills.
PUB DATE Apr 95
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 1995).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academically Gifted; High Schools; *High School Students; Science Education; *Science Process Skills; *Scientific Concepts

ABSTRACT

The selection criteria of high intellectual ability and strong academic achievement for gifted educational programs often leads teachers to assume that their gifted students understand important science concepts. However, empirical evidence of such homogeneity among gifted students' deep understanding of scientific concepts is lacking in the research literature. The objectives of this study were: to describe the level of scientific reasoning ability of high school students in a gifted education program and to examine the viability of using analogous problems and questions designed to measure understanding of basic scientific concepts and skills. The results indicate that the development of valid assessments of science understanding is a key need for both individual student assessment and curricular evaluation. Another finding of the study, that even gifted students are not necessarily equal with respect to their ability to solve different kinds of scientific problems, supports the need for rigorous diagnostic assessment of students' conceptual science understanding and calls for increased small group or independent learning activities in science teaching. Recommendations from this study would be to assess gifted students' preconceptions of scientific concepts and to use multiple measures in judging students' scientific understanding since task-specific effects are very likely. (Contains 30 references.) (JRH)

Scientific Understanding in High Ability High School Students: Concepts and Process Skills

by
George M. Bass, Jr.
and
Roger R. Ries
College of William and Mary

Presented at AERA Annual Meeting
San Francisco, CA
April 18, 1995

BEST COPY AVAILABLE

Scientific Understanding in High Ability High School Students - Concepts and Process Skills

George M. Bass, Jr. & Roger R. Ries
College of William and Mary

Presented at AERA Annual Meeting
San Francisco, CA April 18, 1995

In an effort to reconceptualize science education during the past decade, key national groups of scientists and science educators have formulated recommendations and standards for improving students' science learning (American Association for the Achievement of Science, 1989, 1990; National Assessment of Educational Progress, 1988; Office of Educational Research and Improvement, 1992). Books such as *Science for all Americans* by Rutherford and Ahlgren (1989) have proposed an integrated scientific view based on key scientific concepts underlying specific curricular content. Since these new approaches stress a deeper understanding of essential science concepts, it is important for teachers to know how to assess such understanding. It is especially important to discover any pre-existing conceptions or intuitions about science that students bring to the classroom. It is also critical to know what scientific process skills these students have acquired during their formal science classes.

Many science reforms have clearly been driven by assessment, especially when that assessment indicates American students are not performing up to expectations. The National Assessment of Educational Progress results have raised serious concerns about what high school students know about science. For example, in the 1990 NAEP report findings, only 7% of 17-year-olds could infer relationships and draw conclusions using detailed scientific knowledge. Such findings have also led to much discussion about the best ways to assess student understanding. The American Association for the Advancement of Science has collected numerous papers exploring science assessment in the context of policy issues, curricular reform efforts, instructional impacts, and field-based examples (Kuhn and Malcolm,1991).

Much of this recent examination of science assessment reflects more general calls for new approaches to educational testing and assessment (Perrone, 1991; Berlak, et.al., 1992; Wiggins, 1993; Reynolds, 1994). Reporting on the Secretary of Education's Third Conference on Mathematics and Science Education, the Office of Educational Research and Improvement (1994) has presented conceptual guidelines for designing new assessment systems in science:

- Assessments must be coupled with higher performance standards.
- Assessment systems must measure what we value as opposed to what is easy to measure.

- Assessment should help students learn mathematics and science.
- Assessments must be equitable.
- Every aspect of an assessment system, including its design, should be consistent with its purpose.
- Teachers must be actively involved in reforming assessment and in assessing students.
- New assessments must be open to review and scrutiny.

Indeed, many states such as California, Arizona, Connecticut, New York, Kentucky, Minnesota, and Vermont are in the process of creating alternative assessment systems that try to follow these recommendations (OERI, 1992).

McDermott (1984) has identified many of the characteristics of research on conceptual understanding that likewise influences students' assessment results - nature of the instrument used; degree of interaction between student and examiner; depth of probing; form of data; physical setting; time frame; and goals of examiner. In addition, the crucial relationship of science instruction to various assessment strategies makes it critical that students' competence is assessed on the science content and teaching methods actually taught. Naturally, all these factors make the development of alternative state ssment efforts so difficult to accomplish, especially when students of all abilities must be examined. Will the general, comprehensive assessment strategies used for all students also be valid for high ability students? How do gifted students perform with respect to different tasks to assess scientific conceptual understanding?

Since the pioneering cognitive science research by Brown and Burton (1978) on understanding student procedural "bugs," there has been an increasing recognition of the importance of learner misconceptions on instructional success. Studies in the area of physics by Clement (1983) and McCloskey (1983) among others have documented the impact of novice learner's "naive" theories about various scientific concepts. Discussions by practitioners on how to deal with such misconceptions have also become increasingly noticeable (Berliner, 1987; Griffiths, Thomey, Cooke & Normore, 1988; Gil-Perez & Carracscosa, 1990; Perkins & Blythe, 1994). Recent influential books [*The unschooled mind* by Howard Gardner (1991); *Schools for thought - A science of learning in the classroom* by John Bruer (1993; *Classroom lessons - Integrating cognitive theory and classroom practice* edited by Kate McGilly (1994)] have continued this theme by identifying how strongly students' misinformation and misconceptions affect their later learning. The selection criteria of high intellectual ability and strong academic achievement for gifted education programs often leads teachers to assume that their gifted students understand important science concepts. However, empirical evidence of such homogeneity among gifted students' deep understanding of scientific concepts is lacking in the research literature. To what degree do high ability students also possess many misconceptions about scientific concepts?

The primary objective of this study was to describe the level of scientific reasoning ability of high school students attending the Governor's School for the Gifted in Science and Technology at the College of William and Mary. A second purpose of this investigation was to examine the viability of using analogous problems and questions designed to measure understanding of basic scientific concepts and skills. Specifically, this study attempted to answer the research question, "What is the level of scientific reasoning and understanding among high ability high school students attending a Governor's School for the Gifted in Science and Technology?"

Methods

Subjects

The Governor's School for the Gifted in Science and Technology at the College of William and Mary involves gifted high school rising juniors and seniors from Virginia who have a special aptitude and interest in science. Since 1990 between 150 and 225 students per year have attended a four-week residential summer program in Williamsburg to receive instruction in one of five fields of science: biology, chemistry, geology, physics/astronomy, and computer science.

The subjects for this descriptive study were the high school students attending The Governor's School for the Gifted in Science and Technology during the 1992 and 1993 summer sessions. These students were selected by their individual school systems according to guidelines established by the Virginia Department of Education. Students were to have a strong interest and aptitude in science and to be representative of the gender, racial, and socio-economic makeup of the local school system.

Tasks and Procedures

The research literature on scientific problem solving was reviewed to identify age-appropriate problems that have been used to measure students' understanding of specific scientific principles. A problem in designing an experiment to determine the effect of exercise on heart rate was selected from the 1987 National Assessment of Educational Progress. Three problems that test subjects' understanding of the relationship between force and motion were also selected [a rocket in space problem used by Clement (1983) and an object going over a cliff problem and a running man dropping a ball used by McCloskey (1983)]. As part of the Cognitive Analysis Project, Renner (1979) developed written science problems to assess a person's level of cognitive development. Three problems were selected from his efforts: a problem in designing an experiment to determine the effect of various factors affecting geranium growth, a proportional reasoning problem in comparing shadows of buildings and posts, and a population sampling problem involving frogs in a pond. (A second form of the frog problem was created by substituting different numbers in the original problem, thus requiring the same reasoning but a different arithmetic calculation.) All three of these problems emphasized scientific reasoning with all factual knowledge needed to solve the problem provided in the problem. Multiple-choice questions assessing the recognition of hypotheses and variables in scientific experiments were taken from the Integrated Process Skills Test II (Okey, Wise & Burns, 1982).

During the first week of the 1992 session all students received the frog (population sampling) problem and rocket in space (force and motion) problem. During the last week of the 1992 four week session, half of the students were randomly selected to receive the same problems they had answered on the pretest. The remaining students received two analogous problems to solve - the shadows (proportional reasoning) problem and the object over the cliff (force and motion) problem.

Students at the 1993 session were randomly selected to receive one of two test forms during the first week. Form A contained the same frog problem and rocket problem used in 1992. In addition students were given the heart problem assessing experimental design and the cliff problem measuring their understanding of a falling body. Form B contained the same rocket problem, the alternate frog problem with different numerical values, the geranium problem assessing experimental design, and the falling ball problem. During the last week of Governor School, students received the other form of the test instruments, that is, students completing Form A for the pretest now had Form B for the posttest while students taking pretest Form B received posttest Form A. [All problems and scoring criteria are found in the Appendix.]

Results

Student performance on the problems was independently scored by two graduate students using the scoring guidelines accompanying the published problems. For example, the population sampling problem involving frogs and proportional reasoning shadows problem used a seven-item scale with the proper use of a proportion to solve the problem being categorized in the highest category 7. The heart and geranium problems were likewise scored into 6 or 7 categories according to the accuracy and completeness of the proposed experimental design. The rocket, cliff, and ball problems were classified into various categories based on the path of the moving object drawn by the students.

Analysis of the 1992 data revealed there was much variability in the level of students' reasoning to specific problems (Table 1). The frog problem revealed that formal operational thinking (scores 5-7) was demonstrated by only 45% of students on the pretest and 55% on the posttest. However, using the shadows problem to measure proportional reasoning resulted in over 95% of the posttest students utilizing formal operational thinking. When the frog problem was given to the 1993 students, 44% of them obtained scores in the formal operational thinking range. On the posttest, this increased to almost 60% of the students attaining this same level of understanding.

Approximately 29% of the 1992 Governor School students drew the correct path on the rocket in space problem on the pretest and 27% of the students answering that same item on the posttest had the proper understanding of the effects of force on motion (Table 2). Using the analogous problem of an object falling over a cliff on the posttest revealed almost 64% correct understanding of that concept. The 1993 students showed less understanding of the rocket problem with less than 13% getting it correct. When the same rocket problem was answered on the posttest, 31% drew the correct path. However, only 11% of the students got the problem correct on both the pretest and posttest. While these proportions may seem low for such a select group of high school students, Clement (1983) reported that only 9% of a sample of 150 entering freshman engineering majors solved the rocket in space problem and only 19% of 43 engineering students got it correct after taking a college mechanics course.

The 1993 students' understanding of the effect of force on motion was also tested with the cliff and ball problems (Table 2). Approximately 37% of those taking the pretest cliff problem got it correct while only 18% showed the same understanding on the ball problem. When these students switched problems on the posttest, they achieved 25% correct on the ball problem and 73% correct on the cliff problem. Again, there was only a small number of students who got both problems correct - 14% with the pretest ball and posttest cliff problem and 7% with the pretest cliff and posttest ball problem.

On the two 1993 pretest problems measuring experimental design, almost 67% of the students' answers were classified into the two highest of the six categories on the heart rate problem while only 13% were classified into the highest three categories on the geranium experiment problem. When students switched problems on the posttest, they attained 58% in the top two heart categories and about 76% now were classified in the highest three categories on the geranium problem. While the scoring criteria are not strictly equivalent on the two problems, such differences do show the challenge of using different tasks to measure students' ability to design a scientific experiment.

On the 1993 posttests students also completed either four multiple-choice items on experimental design with sugar water or leaves in soil (Table 4). These recognition items were fairly easy for the students with 74% getting all four leaves in soil items correct and 79% getting all four sugar water items correct. However, when the scores of these same students were compared on the heart and geranium problems respectively, there was much variability in their scores. Being able to recognize concepts in a multiple choice format does not necessarily predict how well you can design an experiment in a more open format.

The performance of Governor School students to these paper and pencil tasks clearly revealed much heterogeneity in their responses on both pretest and posttest problems. There was also a range of adequate and inadequate responses to both the science concept and the process problems. Even students showing a correct conceptual understanding on one problem would not necessarily perform adequately on an analogous problem. The two problems used to assess the students' ability to design valid scientific experiments also revealed a lack of mastery of key experimental design considerations in many of these high ability students. Once more students performed differently on the two problems involving essentially the same design issues. Such large differences strongly support the conclusion that the choice of a particular problem and the scoring criteria are more influential on a student's measured understanding than the identified scientific reasoning hypothesized to solve the problems.

Discussion

Recent calls for science teaching reform appropriate to high ability students have incorporated many current cognitive science learning principles. VanTassel-Baska, et. al. (1992) proposes adapting these science curricular reform recommendations to fit characteristics of gifted learners. Specifically, she identifies the dimensions of content-based mastery, in-depth small group and independent learning opportunities, and multidisciplinary exploration of scientific issues and ideas as important components in any science curriculum development effort. Consistent with such a viewpoint is an in-depth understanding of key science concepts and processes rather than memorization of facts and algorithmic procedures. There has also been an increasing recognition that the conceptions students bring with them to the classroom are an extremely important factor for instructional effectiveness.

A key need for both individual student diagnostic assessment and curricular evaluation is the development of valid assessments of science understanding. Tobin, Kahle-Barry and Fraser (1990) have proposed assessing higher-level cognitive learning by incorporating the four R's of rigor, relevance, representative structure, and rational powers. While these criteria provide a valuable theoretical perspective to the development of science problem-solving skills and tasks, they seem more focused on instructional strategies than student learning measurement.

Shavelson, Baxter, and Pine (1991) and, more recently, Adams and Callahan (1995) have addressed the difficulty of assessing science achievement through more process oriented tasks. The findings from this current study support their concerns that students do not perform equally on science tasks designed to be equivalent measures. Apparently analogous problems are often perceived and answered differently by students. As Lipson (1987) has argued, "anticipation of a test and a test format influences both conscious and unconscious decisions that affect what and how we learn" (p.27). If teachers want students to master fundamental science concepts and skills and be able to apply them in unfamiliar situation, these students must to be exposed to a variety of different assessment strategies that encourage such transfer.

Another finding of this study supports the need for rigorous diagnostic assessment of students' conceptual science understanding. Even gifted students are not necessarily equal with respect to their ability to solve different kinds of scientific problems. Such heterogeneity among this sample of high ability high school students also supports the calls for increased small group or independent learning activities in science teaching.

Therefore, one major recommendation from this descriptive study of Governor's School students would be to assess gifted students' preconceptions of scientific concepts since even in this select sample there is much variability in their understanding. Assuming that all high ability students have already learned essential concepts and principles of science is likely to be a "teacher misconception."

Another recommendation would be to use multiple measures in judging students' scientific understanding since task-specific effects are very likely. This conclusion also supports performance assessment research showing that a substantial number of tasks and assessment methods are needed to get a generalizable measure of a student's understanding of important scientific concepts (Shavelson, Baxter & Pine,1991). If students are expected to construct a deeper understanding of science concepts, then teachers must develop a deeper understanding of cognitive assessment.

REFERENCES

Adams, C. M. and Callahan, C.M. (1995). The reliability and validity of a performance task for evaluating science process skills. *Gifted Child Quarterly, 39,* 14-20.

Berliner, D. (1987). How do we tackle kids' science misconceptions? *Instructor, 97*, 14-15.

Berlak, Harold, et al. (1992) *Toward a new science of educational testing and assessment.* Albany, NY: State University of New York Press.

Bruer, J. (1993). *Schools for thought: A science of learning in the classroom.* Cambridge, MA: The MIT Press.

Clement, J. (1983) A conceptual model discussed by Galileo and used intuitively by physics students. In Gentner, D., & Stevens, A. (1983). *Mental models.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Duschl, R. & Hamilton, R. (1992). *Philosophy of science, cognitive psychology, and educational theory and practice.* Albany, NY: New York Press.

Educational Testing Service (1987). *Assessment in the service of learning.* Princeton, NJ: Author.

Gardner, H. (1991). *The unschooled mind.* New York: Basic Books.

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

Gil-Perez, D. & Carrascosa, J. (1990). What to do about science "misconceptions." *Science Education, 74*, 531-540.

Griffiths, A.K., Thomey, K., Cooke, B. & Normore, G. (1988). Remediation of student-specific misconceptions relating to three science concepts. *Journal of Research in Science Teaching, 25*, 709-719.

Kulm, G., & Malcom, S. (1991). *Science assessment in the service of reform.* Washington, D.C.: American Association for the Advancement of Science.

Lipson, J. I. (1987). Testing in the service of learning science: Learning-assessment systems that promote educational excellence and quality. In *Assessment in the service of learning. (Proceedings of the 1987 ETS Invitational Conference).* Princton, N.J.: Educational Testing Service.

McClosky, M. Naive theories of motion. In Gentner, D., & Stevens, A. (1983). *Mental models.* Hillsdale, NJ: Lawrence Erlbaum Associates.

McDermott, L.C. (July 1984). Research on Conceptual Understanding in Physics. *Physics Today,* 24-32.

McGilly, K. (1994). *Classroom lessons: Integrating cognitive theory and classroom practice.* Cambridge, MA: The MIT Press.

Okey, J.R., Wise, K.C. & Burns, J.C. (1982). *Integrated Process Skills Test II.* Athens, GA: University of Georgia.

Paulu, N. (1994). *Improving math and science assessment.* Washington, D.C.: Office of Educational Research and Improvement U.S. Department of Education.

Perkins, D. & Blythe, T. (1994). Putting understanding up front. *Educational Leadership, 51,* 4-7.

Perrone, V. (1991). *Expanding student assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Renner, J.W. (1979). The relationships between intellectual development and written responses to science questions. *Journal of Research in Science Teaching, 16,* 279-299.

Reynolds, C. (1994) *Cognitive assessment: A multidisciplinary perspective.* New York: Plenum Press.

Romey, W. (1988). *Teaching the gifted & talented in the science classroom.* Washington, D. C: National Education Association.

Rutherford, F.J. & Ahlgren, A. (1989). *Science for All Americans.* Oxford: University Press.

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

Tobin, K., Kahle, J., & Fraser, B. (1990). *Windows into science classrooms: Problems associated with higher-level cognitive learning.* London: Falmer Press.

Tom, A. (1992). *Toward a new science of educational testing and assessment.* Albany, NY: New York Press.

U.S. Department of Education Office of Educational Research and Improvement. (1991). *Improving the math and science curriculum: Choices for state policymakers.* Washington, D.C.: Author.

VanTassel-Baska, J., Bailey, J., Gallagher, S. & Fetting, M. (1992). *A conceptual overview of science education for high ability learners.* Unpublished manuscript, The College of William and Mary, Williamsburg, Virginia.

Wiggins, G.P. (1993). *Assessing students' performance: exploring the purpose and limits of testing.* San Francisco: Jossey-Bass.

Table 1

Proportional Reasoning: Frog Problem / Shadows Problem

1992

Scale Score	N=196 PRETEST (Frog Problem)	N=95 POSTTEST (Frog Problem)	N=101 POSTTEST Only (Shadows Problem)
1	20 (10.2%)	16 (16.8%)	4 (4.0%)
2	19 (9.7%)	11 (11.6%)	0 (0.0%)
3	52 (26.5%)	15 (15.8%)	0 (0.0%)
4	16 (8.2%)	0 (0.0%)	0 (0.0%)
5	30 (15.3%)	7 (7.4%)	8 (7.9%)
6	40 (20.4%)	39 (41.1%)	47 (46.5%)
7 [High]	19 (9.7%)	7 (7.4%)	42 (41.6%)

1993

Scale	Frog Problems		Frog Problems	
	N=107	N=104	N=100	N=105
Score	PRETEST FORM A	PRETEST FORM B	POSTTEST FORM A	POSTTEST FORM B
1	8 (3.8%)	7 (3.3%)	5 (2.4%)	4 (1.9%)
2	4 (2.9%)	10 (4.7%)	7 (3.4%)	3 (1.5%)
3	30 (14.2%)	37 (17.5%)	23 (11.2%)	29 (14.1%)
4	12 (5.7%)	10 (4.7%)	9 (4.4%)	3 (1.5%)
5	8 (3.8%)	12 (5.7%)	23 (11.2%)	23 (11.2%)
6	36 (17.1%)	21 (9.9%)	25 (12.2%)	32 (15.6%)
7 [High]	9 (4.3%)	7 (3.3%)	8 (3.9%)	11 (5.4%)

Table 2

Experimental Design:

1993	Heart Problem	Geranium Problem	Heart Problem	Geranium Problem
Scale	N=107	N=104	N=100	N=104
Score	PRETEST FORM A	PRETEST FORM B	POSTTEST FORM A	POSTTEST FORM B
1	6(5.6%)	0(0%)	5(5.0%)	0(0%)
2	10(9.3%)	8(7.7%)	9(9.0%)	6(5.7%)
3	9(8.4%)	46(44.2%)	14(14.0%)	12(11.4%)
4	12(11.2%)	37(35.6%)	14(14.0%)	7(6.7%)
5	49(45.8%)	8(7.7%)	15(15.0%)	8(7.6%)
6	21(19.6%)	4(3.8%)	43(43.0%)	41(39.0%)
7 [High]		1(1.0%)		31(29.5%)

Table 3

Motion and Force:
1992

Response	Rocket Problem N=196 PRETEST	Rocket Problem N=101 POSTTEST	Cliff Problem N=95 POSTTEST Only
Correct Answer	56 (29%)	27 (27%)	59 (62%)

1993

Scale	Rocket Problem N=98 PRETEST FORM A	Rocket Problem N=94 PRETEST FORM B	Rocket Problem N=94 POSTTEST FORM A	Rocket Problem N=98 POSTTEST FORM B
Correct Answer	12 (12%)	12 (14%)	31 (33%)	29 (29%)
	Cliff Problem	Ball Problem	Cliff Problem	Ball Problem
Correct Answer	36 (37%)	17 (18%)	69 (73%)	24 (25%)
	Both Rocket & Cliff	Both Rocket & Ball	Both Rocket & Cliff	Both Rocket & Ball
Correct Answer	7 (7%)	2 (3%)	24 (25%)	10 (10%)

Rocket problem

% correct on both Pre A & Post B =	11 of 98	(11%)
% correct on both Pre B & Post A =	10 of 94	(11%)

Ball and Cliff problems

% correct on both Pretest B & Posttest A =	13 of 94	(14%)

Cliff and Ball problems

% correct on both Pretest A & Posttest B =	7 of 98	(7%)

Table 4

106 students received the four multiple choice items on experimental design with sugar water
84 got all four correct

104 students received the four multiple choice items on experimental design with leaves in soil
77 got all four correct

Experimental Design: Heart Problem / Geranium Problem
1993

Scale Score	N=77 POSTTEST FORM A	N=84 POSTTEST FORM B
1	2 (2.6%)	0 (00.0%)
2	7 (9.1%)	6(7.1%)
3	10 (13.0%)	9(10.7%)
4	11 (14.3%)	7(8.3%)
5	14 (18.2%)	7(8.3%)
6	33 (42.9%)	29(34.5%)
7 [High]		26(31.0%)