DOCUMENT RESUME

ED 028 089 SE 006 292

By-Nardine, Frank E.

Test Every Senior Project: Evidence of Cognitive Processes Related to Science.

Educational Research Council of America, Cleveland, Ohio.
Pub Date Feb 69
Note-28p.; Paper presented at the National Association for Research in Science Teaching Conference, Pasadena, California, February 1969
EDRS Price MF-$0.25 HC-$1.50
Descriptors-*Academic Achievement, Chemistry, *Cognitive Processes, Cognitive Tests, College Bound Students, Comparative Analysis, *Physical Sciences, Physics, Scientific Concepts, Scientific Enterprise, *Secondary School Science, *Student Characteristics, Student Evaluation
Identifiers-Educational Research Council of America, Processes of Science Test, Test Every Senior Project

Reported is a study designed to evaluate differences in cognitive processes related to science among (1) college bound high school students who had studied both physics and chemistry, (2) college bound students who had not studied either subject, and (3) non-college bound students who had not studied either subject. The test used to assess the degree of understanding of scientific principles and methods of scienc e was the Processes of Science Test (POST). In general, Group 1 subjects demonstrated general competence in areas measured by POST and were superior to students in the other two groups. Some differences in the severity of cognitive deficiencies were obtained between Groups 2 and 3 but, in general, students without course work in physics and chemistry demonstrate the same general cognitive weaknesses whether college bound or not. (GR)

ED028067

TEST EVERY SENIOR PROJECT:

EVIDENCE OF COGNITIVE PROCESSES RELATED TO SCIENCE

Frank E. Nardine

Educational Research Council of America

Cleveland, Ohio

Paper Presented At:

National Association for Research in Science Teaching

Pasadena, California

February 8, 1969

TEST EVERY SENIOR PROJECT:

EVIDENCE OF COGNITIVE PROCESSES RELATED TO SCIENCE[1]

Frank E. Nardine

In recognition of the increasing importance of science in our society, the Science Department of the Educational Research Council has as a prime goal the development of science curriculum for all students -- science oriented, as well as non-science oriented. A major emphasis has been the study of those high school students who choose not to elect physics and chemistry. In part, the Test Every Senior Project was undertaken to ascertain the level of understanding of general scientific principles and reasoning ability possessed by these students.

Currently an experimental curriculum entitled SCIENCE PROBLEMS is being developed in order to foster a greater interest in studying science in those high school students who choose not to elect advanced level science courses. The investigation reported is one aspect of our research conducted to better understand the non-science student in relation to the science oriented one.

During a trial year of the SCIENCE PROBLEMS curriculum materials, the science staff observed that boys rather than girls generally elected to enroll in the course. Accordingly, male high school students were drawn from the

1 The author wishes to thank William V. Mayer and Wimburn L. Wallace for their cooperation in providing unpublished BSCS and POST reports.

Test Every Senior Project data bank (see Gallagher, 1969) and studied in relation to four independent variables: stated intention to attend college (henceforth referred to as college bound), stated intention not to attend college (henceforth referred to as non-college bound), course work in both physics and chemistry, and no course work in either physics or chemistry. The dependent variable was the Processes of Science Test (POST) score.

POST was one of eight scales selected to be administered to approximately 1,500 high school students or 1/8 of the Test Every Senior Project population. POST, known as the Impact Test in its early form, was developed as one phase of the BSCS evaluation program.

The specific purpose of POST was to assess the degree of understanding of scientific principles and methods of science fostered by the BSCS curriculum materials. However, the test developers maintain that POST is useful in a more generic sense. In the POST manual (p. 3) it is stated: "Since the test was specifically prepared to appraise a student's understanding of general scientific principles and scientific reasoning ability, it is also useful for courses other than biology in which understanding of the processes of science is important." The manual further states that: "Although the scientific principles are framed in a setting of biological science, knowledge of biology is not a prerequisite for scoring high on the test." Specifically the concerns of the authors were with:

"the methodology of science"

"the bases for judging facts, principles, and concepts"

"the extent to which the student had developed standards for judging or appraising data"

"the student's ability to interpret qualitative and quantitative data"

"[the student's] ... ability to screen and judge the design of experiments"

Furthermore, the test measures:

"the ability of students to recognize adequate criteria for accepting or rejecting hypotheses"

and finally POST assesses the pupil's ability:

"to evaluate the general structure of experimental design in science, including the need for controls, repeatability, adequate sampling, and careful measurement."

The concerns of the authors of POST are in concert with the concerns and objectives of the Science Department of the Educational Research Council. Knowledge of the kind and degree of understanding measured by POST was expected to be useful as a partial basis for revision and further curriculum planning of SCIENCE PROBLEMS materials.

Specifically, the questions asked in this study were:

(1) Do college bound male high school seniors with course work in physics and chemistry score significantly higher on POST than male coll ge bound seniors without course work in physics or chemistry?

(2) Do college bound male high school seniors with course work in physics and chemistry score significantly higher on POST than non-college bound males without course work in physics or chemistry?

(3) Do college bound male high school seniors without course work in physics or chemistry score significantly higher on POST than non-college bound male high school seniors likewise without course work in physics or chemistry?

(4) Does an inspection of each group's item response with its associated cognitive ability category reveal differences in general understanding?

<u>Subjects</u>

Of the 1,512 high school seniors to whom POST was administered, 742 high school males (see Table 1) comprised the pool from which the subjects for the present investigation were drawn:

Group I consisting of college bound males with course work in physics and chemistry totaled 257 subjects.

Group II consisting of college bound males without course work in either physics or chemistry totaled 121 subjects.

Group III consisting of non-college bound males without course work in either physics or chemistry totaled 86 subjects.

The data collection procedures used in the Test Every Senior Project relied almost exclusively upon student response for obtaining personal data and other vital statistics. This methodology precluded obtaining I.Q. scores, particular course grades, and the like. The criteria for the classification of male subjects into the three groups almost assuredly separated them by ability. However, no direct test was possible to establish this fact. An analysis was made to ascertain what percentage in each group had taken a course in General Biology. The results indicated that:

99% of Group I (college bound students with course work in physics and chemistry) had taken General Biology

82% of Group II (college bound students without course work in either physics or chemistry) had taken General Biology

80% of Group III (non-college bound students without course work in physics or chemistry) had taken General Biology

Thus, a vast majority of subjects in this study elected General Biology as a course of study and making the groups somewhat comparable in this respect.

It is quite evident from Table 1 that the test population of the Test Every Senior Project as represented by subjects to whom POST was administered does not represent a cross section of American high school pupils. For example, 35% of the boys elected both physics and chemistry and 85% indicated intentions to go to college.

Results

The means and standard deviations of scores for the three groups on POST are shown in Table 2. The first three hypotheses advanced considered the relationship of science course work and college intentions. Since the evaluation of these hypotheses consisted of testing for the significance of the difference between Group I and II, Group I and III, and Group II and III. t tests[2] were carried out.

Hypothesis 1 predicted that college bound male high school seniors who have taken high school physics and chemistry (Group I subjects) would score significantly higher on POST than college bound male seniors who have not have either of these courses (Group II subjects). The t value obtained indicated a significant difference in the predicted direction ($t = 9.19$, $df = 378$, $p = < .001$). Group I subjects obtained higher scores than Group II subjects.

Hypothesis 2 predicted that college bound male high school seniors with course work in physics and chemistry would score significantly higher on POST (Group I subjects) than male high school seniors who had not elected either physics or chemistry (Group III subjects). The resultant significant t value indicated that Group I subjects' POST scores were significantly higher than Group III's ($t = 14.56$, $df = 241$, $p < .001$).

---

[2] All p values are two-tailed.

Hypothesis 3 predicted that college bound male high school seniors without course work in physics or chemistry (Group II subjects) would achieve significantly higher POST scores than non-college bound male high school seniors who hadn't taken either physics or chemistry (Group III students). The t value obtained indicated that Group II subjects did achieve significantly higher POST scores than did Group III subjects ($t = 5.05$, $df = 205$, $p < .001$).

Thus, the first three hypotheses were confirmed.

Question 4 asked if an inspection of each group's individual item responses and associated cognitive ability classification revealed differences in general understanding among the three groups. In other words, if the group mean scores were significantly different from one another, an item by item analysis might indicate the command of scientific understanding and reasoning that a particular group had demonstrated by the POST performance. This kind of clinical analysis would have real utility in the classroom in that it offers a useful interpretation and application of the POST results for the educator.

The first step in the analysis of individual test items was to ascertain on what items the responses of the three groups were significantly different from one another. Chi square was employed to test the significance of the difference in response to each item between Group I and II, Group I and III, and Group II and III. The results are shown in Table 3. Responses of Group I subjects were significantly different from Group II subjects on 33 of the 40 POST items. Responses of

Group I subjects were significantly different from Group III subjects on 38 of the 40 POST items. Responses of Group II subjects were significantly different from Group III subjects on 10 of the 40 POST items. Almost without exception, Group I subjects chose proportionately more of the correct item responses than did Group II or Group III subjects. The only reversals occurred on items #6, #9, and #22, and only on item #9 was the difference significant. There were no reversals between Group II and Group III. In every instance, Group II subjects chose a significantly greater percent of correct responses than did Group III subjects. On only one question (#22) were no statistical differences found between any of the three paired groups. Item #22 asked what subject matter training — (1) chemistry only, (2) physics and chemistry only, (3) mathematics and physics only, or (4) chemistry, physics, and mathematics — would assist a biologist to understand blood better. The item analysis data available from the test developers also showed an extremely low correlation between this item and total test score.

The items to which more than 50% of subjects within each group responded incorrectly are presented in Table 4. More wrong-than-right responses were made by Group I subjects on 4 items. More wrong-than-right responses were made by Group II subjects on 16 items. More wrong-than-right responses were made by Group III subjects on 27 items. There were no group reversals. That is, if more than 50% of Group I subjects missed a particular item, subjects in Group II and III did also; if more than half of Group II subjects responded incorrectly to a particular item, Group III subjects did likewise.

In order to discern what cognitive patterns or trends were indicated by the significant differences that existed in item response among the three groups of subjects, the "BSCS Grid for Test Analysis" was utilized. Basically this grid classifies each individual test item according to cognitive processes or ability categories involved in answering the item correctly. Of the four major ability categories detailed only three applied to POST. The headings for these categories are:

A. Ability to recall and reorganize materials learned

B. Ability to apply knowledge to new concrete situations

C. Ability to use skills in understanding scientific problems.

According to the grid classification system, 15 POST items fit most appropriately under Category A, 1 POST item fits most appropriately under Category B, and 24 items fit most appropriately under Category C. The entire grid can be found in Appendix A. A more extensive report of the BSCS grid and its development has been written by Klinckmann (1963).

The specific cognitive processes are indicated by the subcategories under each major category heading. The subcategories and the distribution of individual POST items are shown in Table 5.

The criterion of more wrong-than-right response was used as an indication of deficiency. According to this criterion Group I did not show a deficiency in any single category. None of the four greater-than-50% - incorrect response items came from the same category, and as with item #22, evidence suggests that items #9, 19, and 20 had low correlation with total test score. Thus these four items provided little discrimination between groups.

Those items on which more than 50% of Group II and III subjects chose the incorrect response are shown in Table 6 and Table 7 respectively. In regard to Group II, it was found that more than half of the group missed:

2 out of 3 items (or 67%) involving screening hypotheses;

4 out of 7 items (or 57%) involving methodology;

2 out of 4 items (or 50%) involving interpreting qualitative data;

2 out of 4 items (or 50%) involving interpreting quantitative data;

2 out of 4 items (or 50%) involving understanding relevance of data to the problem;

3 out of 8 items (or 38%) involving screening and judging design of experiments.

Thus, a majority of subjects displayed deficiency in six of the POST categories.

The deficiencies as indicated by Group III's performance on POST are summarized in Table 7. More than half of Group III subjects missed:

4 out of 4 items (or 100%) involving interpreting qualitative data;

6 out of 7 items (or 86%) involving methodology;

3 out of 4 items (or 75%) involving understanding relevance of data to the problem;

2 out of 3 items (or 67%) involving screening hypotheses;

5 out of 8 items (or 62%) involving screening and judging design of experiments;

2 out of 4 items (or 50%) involving interpreting quantitative data;

3 out of 8 items (or 38%) involving criteria.

Thus, a majority of subjects displayed deficiency in seven of the POST categories.

Ability categories B1 and C7 having to do with "non-quantitative knowledge" and "analyzing scientific reports" respectively were not included in the summaries of Group II and III as POST contains only one item for each category. This provision of a single item in a category was insufficient to allow meaningful interpretation.

DISCUSSION

The classification of actual POST items (Table 5) according to the BSCS Test Grid ability categories revealed a disproportionate number of items among the several categories. Category A2-4, (criteria) and C3 (screen and judge design of experiments) are represented by 8 items each, while C4 (screen hypotheses) is represented by ? items and B1 (non-quantitative knowledge) and C7 (analyze scientific reports) have but 1 test item each. This imbalance is regretable despite Klinckmann's (1963) explanation that items were classified according to "maximum ability required" by the student. Her implication is that item imbalance is not present. In any case, if a single test item involves several abilities or cognitive processes, then it would be valuable to so specify as precisely as possible. The interested educator with his personal knowledge of the student would then be in a better position to evaluate the individual's performance.

An attempt was made to have the professional staff of the Science Department of the Educational Research Council match items and categories. A great deal of disagreement resulted. The reason for the disagreement was not that items could be multi-classified, thus making classification difficult. Rather, the problem encountered was that the category description of science processes was complex and highly sophisticated in comparison to specific test items. A more careful and detailed explanation of the categorization of each item in POST might have reduced disagreement.

Anyone who has been involved in test construction, and, specifically, item writing can attest to the fact that it is extremely difficult to classify the cognitive processes involved in learning. There is always the hazard of claiming to measure separate or distinct cognitive processes based on convention rather than fact. Yet despite the complexity and admitted failings of isolating and categorizing thought processes, the attempt of the POST authors to develop a test according to a cognitive schema is commendable. At the present time educators spend millions of dollars annually on testing programs without comparable benefits to the teaching-learning process accruing. Generally a composite test score merely indicates to the teacher that a student has mastered or failed to master the science material. Seldom do test results portray students' strengths and weaknesses and so reveal areas in which buttressing or shoring up of teaching effort can and should be concentrated. POST makes a real contribution to science education by providing a framework by which a student's composite numerical score can be broken down and analyzed in cognitive terms which have diagnostic value and implications. The BSCS Grid can serve as a valuable approach to test interpretation, for as Tyler (1968, p. 63) maintains: "...when we are trying to appraise the curriculum by ascertaining how well the students are learning, we need information about those things that nearly all students are learning and those that very few are learning as well as those that are being learned by about half the class." Further development and refinement of the BSCS Grid promises to make consequent testing a more meaningful tool for evaluation of science learning and understanding.

Certain conclusions can be drawn based on the findings of this investigation which may provide a basis for understanding the results of various science curriculums and for planning and developing new educational programs. The usual cautions against overgeneralization must apply, however, due to the restricted sample, all male subjects and the like. Specifically, it is possible to do a between group comparison as well as a within group comparison and arrive at an approximation of the student's grasp of certain cognitive processes. To illustrate, Group I subjects, college bound with course work in both physics and chemistry, demonstrated general competence in the areas measured by POST. As would be expected, these students who were motivated to enroll in the "difficult" courses and who had received all the science training generally offered to high school pupils, demonstrated relative superiority to students who had for some reason not enrolled in physics and chemistry.

Group II subjects, college bound but without course work in either physics or chemistry, demonstrated competence in the areas of "criteria" (A2-4) and "screening and judging design of experiments" (C3).

Group III subjects, non-college bound and without course work in physics or chemistry, showed some degree of competence only in the area of "criteria" (A2-4).

In comparing Groups II and III, it is possible to make determinations about the severity of cognitive deficiencies. For example, in terms of rank ordering Group II is most deficient in "ability to screen hypotheses," whereas Group III's greatest deficiency is "ability to interpret qualitative data." The relative lack of significant differences between Group II and III indicates that, by and large, students without course work in physics and chemistry demonstrate the same general cognitive weaknesses or deficits whether from a college bound sample or not. With the Test Every Senior Project data bank, it will be possible to ascertain whether course work in physics or in chemistry contributes more toward a higher POST score, or whether both must be taken to overcome the cognitive lacks shown by Group II and III students.

Benjamin Bloom (1968) has stated that within the various subject matter areas there ought to be a clear and definite specification of the expected level of learning which indicates mastery. The finding that over 50% of students in both Groups II and III were not able to answer correctly even half of the items in a high percentage of categories demonstrates the necessity of following Bloom's advice. We need to press the enquiry into the skills and understandings which students are or are not learning. Furthermore, a variety of science courses needs to be developed to attract the non-science oriented high school student so that he will be exposed to the type of mental training a study of science offers. Educators and curriculum developers must approach the teaching-learning process in science subject areas with an eye toward developing competence in the various processes of science which contribute to general overall mastery.

Likewise, those engaged in the development of testing and evaluation of programs and those involved in research in science education must continue to develop and refine instruments which will yield a meaningful chart or profile of a student's progress in mastering the various aspects of the processes of science.

Acknowledgement

The research reported in this paper was supported by the Educational Research Council of America, Cleveland, Ohio, and by a grant from the Sears Roebuck Foundation, Skokie, Illinois.

TABLE 1.

PHYSICS AND/OR CHEMISTRY COURSE WORK AND COLLEGE

INTENTIONS OF 742 MALE HIGH SCHOOL SENIORS.

	Physics & Chemistry	Physics	Chemistry	No Physics or Chemistry	Other*	Total
College bound	257	32	171	121	47	628
Non-college bound	5	4	15	86	4	114

* Incomplete data prevented classification

TABLE 2.

MEANS AND STANDARD DEVIATIONS OF SCORES ON

PROCESSES OF SCIENCE TEST.

	N	M	SD
Group I	257	28.58	5.95
Group II	121	22.00	6.79
Group III	86	17.40	6.22

TABLE 3.

ITEMS ON WHICH CORRECT RESPONSES WERE SIGNIFICANTLY DIFFERENT FROM ONE ANOTHER BETWEEN THE PAIRED GROUPS IN THE DIRECTION INDICATED.

Item	GROUPS I > II*	I > III	II > III
1		Z	Y
2		X	
3	X	Z	Y
4	Z	Z	
5	Y	Z	
6		Z	X
7	Y	Z	
8	Z	Z	
9	Z*		
10	X	Z	
11	Z	Z	
12	Z	Z	
13	Z	Z	Y
14	Z	Z	X
15	Z	Z	
16	Y	Z	X
17	Z	Z	
18	Z	Z	
19		X	
20		Z	
21	Z	Z	
22			
23	Z	Z	
24	Y	Z	
25		Y	
26	Z	Z	
27	Z	Z	
28	Y	X	Z
29	Z	Y	X
30	Y	Y	
31	Y	Y	
32	Z	X	Y
33	Z	Z	
34	Y	Z	Z
35	Y	Z	
36	Z	Z	
37	Z	Z	
38	Z	Z	
39	Z	Z	
40	Y	Z	

X indicates p = > .05 Y indicates p = > .01 Z indicates p = > .001

* On Item #9 Group II subjects chose significantly more correct responses in proportion to Group I subjects.

TABLE 4.

ITEMS ON WHICH MORE THAN 50% OF SUBJECTS WITHIN A

GROUP RESPONDED INCORRECTLY.

	GROUP		
Item	I	II	III
1			X
2			
3			X
4			
5			
6			X
7			
8			X
9	X	X	X
10			X
11			X
12		X	X
13		X	X
14		X	X
15			X
16			
17		X	X
18			X
19	X	X	X
20	X	X	X
21		X	X
22	X	X	X
23			
24		X	X
25		X	X
26		X	X
27			
28			X
29			X
30		X	X
31			
32			
33			
34			X
35			
36		X	X
37			
38		X	X
39		X	X
40			

TABLE 5.

CLASSIFICATION OF POST ITEMS ACCORDING TO

BSCS TEST GRID ABILITY CATEGORIES.

Ability Category			POST ITEM NUMBER
	A2-4	Criteria	3, 29, 31, 32, 33, 34, 35, 40
	A2-5	Methodology	2, 11, 20, 21, 26, 28, 30
	B1	Non-quantitative Knowledge	22
	C1-1	Interpret qualatitive Data	10, 15, 36, 38
	C1-2	Interpret quantitative Data	4, 5, 12, 13
	C2	Understand relevance of data to problem	9, 18, 24, 37
	C3	Screen and judge design of experiment	1, 6, 7, 14, 16, 17, 19, 27
	C4	Screen hypotheses	23, 25, 39
	C7	Analyze scientific reports	8

TABLE 6.

POST ITEMS CLASSIFIED BY ABILITY CATEGORY ON WHICH

MORE THAN 50% OF GROUP II SUBJECTS RESPONDED INCORRECTLY.

Ability Category	A2-4 (8)*	A2-5 (7)	B-1 (1)	C1-1 (4)	C1-2 (4)	C2 (4)	C3 (8)	C4 (3)	C7 (1)
POST Items		#20 #21 #26 #30	#22	#36 #38	#12 #13	#9 #24	#14 #17 #19	#25 #39	
Totals	0	4	1	2	2	2	3	2	0

* Maximum number of items within the category

TABLE 7.

POST ITEMS CLASSIFIED BY ABILITY CATEGORY ON WHICH MORE THAN 50% OF GROUP III SUBJECTS RESPONDED INCORRECTLY.

Ability Category	A2-4 (8)*	A2-5 (7)	B-1 (1)	C1-1 (4)	C1-2 (4)	C2 (4)	C3 (8)	C4 (3)	C7 (1)
POST Items	#3 #29 #34	#11 #20 #21 #26 #28 #30	#22	#10 #15 #36 #38	#12 #13	#9 #18 #24	#1 #6 #14 #17 #19	#25 #39	#8
Totals	3	6	1	4	2	3	5	2	1

* Maximum number of items within the category

APPENDIX A

BSCS TEST GRID CATEGORIES[1]

---

[1] Taken from BSCS NEWSLETTER, #19, September, 1963, 18-19.

BSCS Test Grid Categories

A. ABILITY TO RECALL & REORGANIZE MATERIALS LEARNED.[1] This category primarily involves remembering—either by recognition or recall—the information which has previously been studied. It also may involve a simple problem, the key to which is being able to recall all the information or details of meaning of certain key terms in the problem. Items falling into this category can have a high degree of difficulty either if they require remembering a considerable amount of information or if they require remembering a complicated kind of information such as a complex theory.

A1. **Memory of specifics.**

A1-1. **Terminology.** Memory of the referents for specific symbols. Examples in biology are cilia, nucleus, petiole, etc. This category may also include memory of the most generally accepted symbol referent, or the variety of symbols which may be used for a single referent, or the referent which is most appropriate to a given use of a symbol.

A1-2. **Specific facts.** Events, persons, dates, locations, descriptive characteristics, etc. This may include very precise information—such as the size of a human red blood cell—as well as approximate information—such as the general order of magnitude or relative size of different kinds of organisms.

A2. **Memory of ways and means of dealing with specifics.** This refers to knowing the ways of studying, organizing, judging and criticizing which are characteristic of biological science. It includes such means as the ordering of sequences, methods of investigation and standards of judgment. These are abstractions intermediate between specific information and universals. This category does not demand that a student be able to use these materials, but only that he have a passive awareness of their nature

A2-1. **Conventions, conceptual mod' and heuristics.** Memory of ways of treating and presenting ideas and phenomena characteristic of biology. These are the usages, styles, practices and forms which best suit the purposes of the workers in the field or which seem to best suit the phenomena which they study. They may be arbitrary, accidental or authoritive, but often are retained because of general agreement of individuals working in the area. An example would be a model of the DNA molecule.

A2-2. **Trends and sequences.** Processes, movements and directions of phenomena with respect to time. Examples are embryological development, mitosis.

A2-3. **Classification and categories.** The classes, sets and arrangements which are regarded as fundamental to a given purpose, argument, subject or problem. E.g., phylogenetic categories, germ layers, functional cell types.

A2-4. **Criteria.** The bases for judging facts, principles, conduct. E.g., Darwin's criterion for distinguishing more variable species from less variable species.

A2-5. **Methodology.** Methods of investigation, inquiry, techniques, technological devices and procedures used in the field of biology as well as employed in the studying of particular problems or phenomena. E.g., the plot technique of ecology, the organ-function approach in physiology, special technological devices for studying living cells.

[1] The sub-categories of A. are same, for the most part, as "Memory" sub-categories in Benjamin S. Bloom, *op. cit.*, pp. 201-207. Explication and examples pertinent to biology have been added.

Again, it should be stressed that this category refers to familiarity with the method or technique, not ability to use it.

A3. **Memory of universals and abstractions in a field.** These are the major concepts, schemes and patterns by which ideas and phenomena are organized. These are at the highest level of abstraction and include the theories and structures which dominate a field and are generally used in its study.

A3-1. **Principles, generalizations and concepts.** The more particular abstractions which summarize a group of phenomena. E.g., principles of enzyme activity, concept of the gene, concept of natural selection.

A3-2. **Theories, structures and conceptual schemes.** A group of principles and generalizations along with their interrelations which constitute a rounded, systematic view of a complex phenomenon, problem or field. These can be used to interrelate and organize a great range of specifics. E.g., evolutionary theory and other BSCS themes.

B. ABILITY TO APPLY KNOWLEDGE TO NEW CONCRETE SITUATIONS. This entails the ability to use remembered knowledge in a new, unfamiliar or fictional situation.

B1. **Application of non-quantitative knowledge.** Application of abstractions to particular, concrete situations. E.g., principles of food webs applied to new, unfamiliar or fictional biotic communities. Principles at the level of abstraction of A 2 and A 3 are likely to be used in this way.

B2. **Application of quantitative materials.** Similar to B1 but the materials applied are quantitative. E.g., use of quantitative data, application of principles of graphing, ratios, etc.

C. ABILITY TO USE SKILLS INVOLVED IN UNDERSTANDING SCIENTIFIC PROBLEMS. (One way to test both the laboratory materials and these abilities is to develop items which parallel, but are not identical to, the investigatory laboratory exercises. The test items could present part of such an experience and invite the student to complete it.)

C1-1. **Interpret qualitative data.** This may include the following:

a. The ability to translate, or paraphrase, a statement of data.

b. The ability to explain or summarize a statement of data. This involves some sort of reordering or new view of the material. It may include the ability to draw the conclusion which fits the problem when presented with the description of an experiment and the data found. A variation on this is to describe an experiment and the conclusion drawn, then ask what sorts of data must have been found to justify such a conclusion.

c. The ability to extend a known principle to account for data which cannot be accounted for by the known principle in its original form.

C1-2. **Intepret quantitative data (graphs, tables, charts).** This may include one or more of the following:

a. The ability to translate, or provide a verbal description of, a table or chart; conversion from table to graph, etc.

b. The ability to explain or summarize graphs, tables, etc. This may entail ability to discern significance of data or to make a choice between conclusions which are more or less justified regarding the data presented.

c. Interpolation and extrapolation. Interpolation is the ability to derive particular quantitative relations within a given series of observed particulars. Extrapolation is the ability to derive a particular relation by extension of the observed series. This includes an understanding of the relative validity of extrapolation and interpolation.

d. Discerning connections and interactions between elements and parts. A number of kinds of relationships are used in biology. For quantitative data these include: expression of graphic data in a mathematical formula and deriving equations from particular sets of data; comparing particular sets of data with a generalization, e.g., chi-square; testing the fit of data to formulae or equations.

C2. **Understand relevance of data to problem.** This may include:

a. The ability to discern different degrees of adequacy of data relative to a stated problem.

b. The ability to understand "best" and "second-best" data, and the reasons why the latter are sometimes the only data that can be obtained.

C3. **Screen and judge design of experiments.** Designing experiments involves development of a plan of work. It is obvious that development of such plans is not required in objective test items. Rather, a recognition of these plans or proposals is all that is required. Such recognition includes understanding of an experiment's oppropriateness relative to a stated problem. Recognition of adequate experimental design may also require an understanding of the concepts of control and adequacy of sample. It may involve discerning the kind of data which can be obtained from a given experiment.

C4. **Screen hypotheses.** This involves the ability to select one hypothesis from several. Selection of the most appropriate hypothesis may be in terms of relevance to the problem, to design of experiment, to data collected, to feasibility of collecting data, etc.

C5. **Identify problems and unanswered questions.** This may include:

a. Ability to reorganize information may result in identification of a new problem or question.

b. Ability to discern inconsistencies and/or logical inaccuracies in known information may suggest a new problem.

c. Ability to discern biological problems growing out of certain areas of relationships in other disciplines, such as physics, mathematics or geology.

C6. **Identify assumptions and principles of inquiry and extend their application and scope.** This includes the following abilities:

a. Identification of the organizational principles of a scientific report, i.e., analysis of the systematic arrangement and logical structure which holds a report together.

b. Application of a principle of inquiry, or a set of principles, to a new or unfamilar research problem.

c. Discerning of a set of assumptions or principles of inquiry which account for the point of view or organizational structure of a report. This includes implicit assumptions and principles of inquiry of which the writer of the report may not be explicitly aware.

d. Extension of the scope of principles of inquiry. This entails modification of a known principle of inquiry and provides a new one which is a more adequate basis for approaching new problems. For example, regulation can be considered an extension of the principle of homeostasis.

C7. **Analyze scientific reports.** This includes:

a. Identification of the elements included in a research paper, e.g., the problem with which it deals, the experimental design, etc.

b. Analysis of the relationships between these elements and parts.

c. Analysis of pervasive organizational principles of scientific reports.

d. Evaluation of scientific reports in terms of internal evidence, i.e., judgment or evaluation of a report on such grounds as logical accuracy and consistency.

e. Evaluation of scientific reports in terms of external criteria, i.e., judgment or evaluation of a report on grounds external to the report itself. The student might be expected to select from a number of sets of criteria that set most appropriate to the report in question; or he may be asked to remember criteria previously learned.

D. ABILITY TO SHOW RELATIONSHIPS BETWEEN BODIES OF KNOWLEDGE. This refers to relating different bodies of knowledge learned at different times or in connection with different topics and requires relating the given bodies of knowledge in ways other than the ways in which they were related in text, lab or class work.[1] The ability sub-categories of this category are defined in terms of different kinds of ways of relating the content material of biological knowledge:

D1. **Comparison.** This ability requires more than simply recalling and contrasting previously learned information, as described in Category A. It requires being able to compare and contrast materials on points not previously learned. This category also includes discerning previously unknown relationships.

D2. **Extrapolation.** This is the extension of trends, tendencies or generalizations beyond the given data or phenomena to determine implications, consequences, effects, etc., which are in accordance with the trend, tendency or generalization in question. This category includes extrapolation in the broad sense of being able to discern, for example, the significance of a theory to future research rather than extrapolation in the more limited sense of extension of trends expressed quantitatively, e.g., prediction of U.S. population 1970. The latter would be classified as C1-2.

D3-1. **Application to another biological area.** Application of concepts or models from one area of biology to phenomena of another area. One example is the notion of a "molecular community." Principles at the level of abstraction of A3 are most likely to be used in this way.

D3-2. **Application to other fields.** For example, the application of ecological principles to certain human social problems.

D4. **Analysis of relationships.** Discerning new connections and interactions between elements and parts. For example, discerning that there may be a relatively constant ratio between nuclear and cytoplasmic materials in different kinds of cells if this has not been previously pointed out in the text or class discussion. May include understanding the logical relations commonly called induction, deduction, analogy, and the differences in their looseness and precision.

D5. **Interrelate facts, principles, phenomena, etc., in a new way.** Discerning and evaluating a new way of organizing or relating specifics within the field biology. E.g., the possible relationship between DNA replication and certain virus-bacteria relationships.

D6. **Development of a set of abstract relations.** The imaginative development of a new set of interrelated concepts.

---

[1]It is very difficult to construct multiple-choice items which test for abilities in this category. It may be that only short essay items can adequately test abilities described in D4 and D3.

REFERENCES

1. Bloom, S. Learning for Mastery. Evaluation Comment. Center for the Study of Evaluation of Instruction Programs. University of California, Los Angeles, May, 1968. Vol. 1, No. 2

2. Gallagher, J. J. Test Every Senior Project: Design and Preliminary Analyses Paper presented by National Association for Research in Science Teaching, 42nd Annual Meeting. Pasadena, California, 1969.

3. Klinckmann, E. The BSCS Grid for Test Analysis. BSCS Newsletter, #19, September, 1963, 17-21.

4. Tyler, R. W. Analysis of Strengths and Weaknesses in Current Research in Science Education. Journal of Research in Science Teaching, Vol. 5, Issue 1, 1968, 52-63.