R E P O R T R E S U M E S

ED 017 467 SE 004 231

IMPROVING OBJECTIVE TESTS IN SCIENCE.

BY- NELSON, CLARENCE H.
NATIONAL SCIENCE TEACHERS ASSN., WASHINGTON, D.C.

PUB DATE 67
EDRS PRICE MF-$0.25 HC NOT AVAILABLE FROM EDRS. 27P.

DESCRIPTORS- *EVALUATION, *TESTING, *TEST CONSTRUCTION, ACHIEVEMENT TESTS, COLLEGE SCIENCE, ELEMENTARY SCHOOL SCIENCE, EDUCATIONAL OBJECTIVES, OBJECTIVE TESTS, PROBLEM SOLVING, SECONDARY SCHOOL SCIENCE, NATIONAL SCIENCE TEACHERS ASSOCIATION,

DISCUSSED ARE SELECTED FACTORS RELATED TO IMPROVING SCIENCE TESTS. SECTION 1 DEALS WITH (1) THE ROLE OF OBJECTIVES IN SCIENCE TEACHING AND TESTING, (2) THE ROLE OF TESTS AND TESTING, (3) FORMS AND KINDS OF TESTS, AND (4) GUIDEPOSTS FOR PREPARING TESTS. SECTION 2 DISCUSSES METHODS OF TESTING FOR SPECIFIC OBJECTIVES AT VARIOUS GRADE LEVELS. INCLUDED ARE (1) RECOGNIZING AND APPRAISING ASSUMPTIONS, (2) UNDERSTANDING OF SCIENTIFIC METHODOLOGY INCLUDING THE FORMULATION AND TESTING OF HYPOTHESES, DATA ANALYSIS, AND THEORY COMPARISON, AND (3) SELECTING AND PREPARING READING PASSAGES FOR FIRST GRADE, THIRD GRADE, UPPER ELEMENTARY GRADES, JUNIOR HIGH SCHOOL LEVEL, AND SECONDARY SCHOOL LEVEL. THIS DOCUMENT IS ALSO AVAILABLE FOR $1.00 FROM NEA PUBLICATIONS SALES, 1201 SIXTEENTH STREET, N.W., WASHINGTON, D.C. 20036. (DS)

Revised Edition of
Let's Build Quality into Our Science Tests

IMPROVING OBJECTIVE TESTS IN SCIENCE

Clarence H. Nelson
Office of Evaluation Services
Michigan State University

National Science Teachers Association, Washington, D.C.

THE NATURE AND STRUCTURE OF SCIENCE

The philosophy of the newer science curricula is that science consists not only of facts, concepts, and principles to be memorized, but also of investigative processes that have been and are being utilized to create new knowledge. Consequently, much greater emphasis is now being placed upon the investigative and inquiry aspects of science, some facets of which are observation, formulation of hypotheses, experimentation, interpretation and analysis of data, and evaluation of their pertinence and relevance to the solution of the problem or problems under consideration.

If every student in the class is given many opportunities to experience scientific activity at a level that he or she can understand, the recurrent complaint of boredom from passive observation will probably be considerably reduced. All scientific activity need not necessarily entail physical moving or handling of apparatus. Use of apparatus can and perhaps should be reduced to a minimum lest this activity become the end in itself. Under optimum conditions the intellectual activity associated with the experiment can prove to be more stimulating for many students than the mere physical handling of the apparatus. This can come about through skillful teaching and the use of laboratory guides that have been written so as to elicit patterns of clear stepwise thinking in the solution of problem situations.

Stimulating rather than perfunctory teaching by well-trained, interested individuals, who themselves have a grasp and appreciation of the nature of science from firsthand science research experience, can do much to counteract the negative image of scientists and scientific work held by a surprisingly large number of high school students. (14) Evaluation of the outcome of such teaching becomes much more meaningful in character and perspective than is possible in a situation where the students have done little more than memorize and echo detailed facts, laws, and principles from the textbook.

As we consider the unique nature and structure of science, we may gain better insights into the procedure for creating the objective science test.

Revised Edition of
Let's Build Quality into Our Science Tests

IMPROVING OBJECTIVE TESTS IN SCIENCE

Clarence H. Nelson
Office of Evaluation Services
Michigan State University

National Science Teachers Association, Washington, D.C.

1201 Sixteenth St., N.W.
Washington, D.C. 20036

Stock Number 471-14432 Single copy, $1

Order copies from NEA Publications Sales
1201 Sixteenth St., N.W., Washington, D.C. 20036
Discounts on multiple orders: 2-9 copies, 10 percent; 10 or more copies, 20 percent
Payment must accompany orders of $2 or less

THE USE OF TESTS

I. THE ROLE OF OBJECTIVES IN SCIENCE TEACHING AND TESTING

How do science classroom teachers know whether the goals of instruction are being met? How do the students and their parents know whether satisfactory progress is being made in the science courses? Why should it be necessary to know?

Course objectives, goals, or outcomes are designed to give direction to all aspects of the course, including course design, teaching techniques, instructional materials, equipment, and the evaluation procedures that will be employed.

After the objectives have been formulated, all the course activities should be planned in such a way that the objectives can be implemented with some prospect that they will be attained. Evaluation of stated objectives should be of prime consideration, for students tend to study and learn those things upon which their achievement is judged.

If objectives are to be useful, it is imperative that the teacher translate general objectives into behavioral objectives that can be attained and measured. It is not enough, for example, to have in mind the teaching of critical thinking, if it is not also clear what specific design will lead to the achievement of this broader objective.

There are many sources which elaborate the behavioral objectives of science teaching, and science teachers should become familiar with them. Those who desire to study objectives at length are referred to the work of Kessen (11).[1]

[1] The following excellent book discusses the methods appropriate to the teaching of science as a human endeavor:
Woodburn, J. H., and Obourn, E. S. *Teaching the Pursuit of Science*. Macmillan Company, New York. 1965.

II. THE ROLE OF TESTS AND TESTING

Foundational to performance of most vocations, and indeed to being a successful participant in our society, is the satisfactory completion of a basic education which usually includes science. To succeed in the professions a very high degree of excellence in the educational foundations becomes a prerequisite, for high proficiency at the more advanced levels becomes almost impossible unless the early groundwork has been well laid. To a certain extent the same is true in other vocations as well. Each occupation has its own standards of excellence. The highly skilled craftsman must be able to make very precise measurements and must acquire a marked degree of sensitivity to quality workmanship. These often correlate highly with skills and habits acquired very early in the classroom where the teacher has inculcated a wholesome respect for standards of excellence. Proficiency in the skills and acquisition of the habits included in the standards were amenable to fairly precise measurement and appraisal. Though testing is only one method of evaluation, it is an important and useful procedure, worthy of the most careful preparation and refinement.

Despite the criticism in some quarters of our competitive system of standards, we probably would not, even if we could, want to abolish the measures of proficiency that our society has established. Our problem, then, is to determine how to improve measurement so as to reflect an acceptable degree of validity. In short, how can we construct tests that will measure rigorously and accurately the attainment of the stipulated outcomes of instruction, and thus be assured of valuable feedback to serve as a guide in systematically structuring the educational program from the earliest years into the collegiate level?

To many teachers, professors, and critics, objective testing implies measurement of factual recall and nothing more. This is indeed true of many objective tests. The author's analysis of final examinations given by a number of classroom teachers revealed that approximately 90 percent of the items required nothing more than repetition of memorized isolated bits of information bearing little relationship to each other or to a central idea. Perhaps the chief reason is that factual recall-type items take least time and effort to construct. It is also likely, however, that most teachers have not had sufficient training in test construction to have mastered alternative procedures that would measure other types of learning from instruction. Elsewhere in this book are example questions to assist such teachers in constructing items to measure comprehension. These need not be used verbatim. They are included primarily to convey style rather than content.

It is possible to create sets of objective test exercises that are distinctive and elegant, whose items are interrelated and unfold a solution to a problem situation.

III. FORMS AND KINDS OF TESTS

Tests may vary in form according to the preference of the teacher who makes them. Some prefer what is commonly known as the essay test, while others tend to favor objective tests. Usually less time is required in the actual construction of an essay test but more time is required to mark the papers afterward. With an objective test the greater time goes into construction of the test. It also requires much thinking, rethinking, and reworking to avoid ambiguity, superficiality, artificiality, and to incorporate questions that inspire thinking rather than mere memorization. Hence, this publication is directed chiefly to the improvement of the objective test in science.

Forms of Tests

1. *Unstructured*—essay, completion. Sometimes called free-response test. The test-taker either supplies or creates the answers.
2. *Structured*—multiple-choice, matching, true-false, etc. The test-taker selects the answer from among alternatives provided.

Uses of Tests

1. *Diagnostic*—sometimes called analytical. Used primarily for evaluating the general level of competence and knowledge of the student and for checking day-to-day progress. Diagnostic tests are *not* primarily for grading purposes. Diagnostic testing is often followed by remedial instruction over those aspects of the course which the students failed to master satisfactorily. It may also be used as a guide for placement of students in various sections or at a particular educational level, and should certainly be used as a guide to the teacher in his plans for a particular course. The expected level of performance on diagnostic tests—at least on those which check day-to-day progress—is high, perhaps in the 80 to 100 percent range.
2. *Achievement*—Given at intervals in the term, primarily for the purpose of assistance in assigning marks. The achievement test usually consists of large blocks, units, or related problems covering a major area. This should be a highly discriminating test with an expected *average* performance of about 50 percent to 60 percent, and with a range of performance extending perhaps from 20 percent to 100 percent.

Some Characteristics of Good Achievement Tests

Validity—The test should measure what it is supposed to measure. This is usually determined by the course content and the course objectives. If the test is valid, the materials embodied within it do not appear strange or foreign to an individual who has attained a degree of competence in the area.

Reliability—The test should measure a person's abilities consistently or in a constant manner, as a steel tape measure, and not as an elastic one. A reliable test should yield very similar results in a series of consecutive measurements. Reliability is largely a function of the clarity of the questions and the number of scorable units in the test. The more items or questions in the test, usually the more reliable it is, providing it is otherwise well constructed.

Objectivity—This increases reliability. A test should receive the same score or grade if it is scored by several different readers or if it is read or marked by the same reader on successive days.

Economy—Reasonable use of the teacher's time, particularly in scoring the papers and interpreting the results.

Fairness—The students perform reasonable tasks in reasonable time allotments.

Motivation—Students can make their best efforts without guessing or bluffing. The foils or distractors in each item should appear so plausible that the "opportunist" student who has not spent adequate time in preparation will be thoroughly confused. However, questions to ensnare all students should not be used. The test should enable the well-prepared student to leave the testing session with a feeling of accomplishment.

Selectivity—Every question or item should separate some students from others; that is, every item should be answered correctly by some students and missed by others.

IV. GUIDEPOSTS FOR PREPARATION OF TESTS

Many test-writers have found it helpful to use the following educational objectives outlined by Bloom and others [1] as a guide to areas to be included in test questions.

Cognitive Domain

Knowledge

- Knowledge of specifics
  - Terminology
  - Specific facts
- Knowledge of use of specifics
  - Conventions
  - Trends and sequences
  - Classifications and categories
  - Criteria
  - Methodology
- Knowledge of the universals and abstractions in a field
  - Principles and generalizations
  - Theories and structures

Intellectual Abilities and Skills

Comprehension—ability to make use of materials or ideas

- Translation
- Interpretation
- Extrapolation

Application—the use of the abstractions in particular and concrete situations

Analysis—making clear the relative hierarchy of ideas in a communication

- Analysis of elements
- Analysis of relationships
- Analysis of organizational principles

Synthesis—the putting together of elements and parts so as to form a whole

- Production of a unique communication
- Production of a plan or proposed set of operations
- Derivation of a set of abstract relations

Evaluation

- By internal evidence
- By external criteria

The Affective Domain

Receiving (attending)
Responding
Valuing
Organization (conceptualization and organization of a value system)
Characterization by value or value complex

[1] Bloom, Benjamin S., and others. *Taxonomy of Educational Objectives*. David McKay Co., Inc., New York. 1956.

A condensed version of the Bloom objectives constitutes one axis of the chart of specifications shown in Figure I, for a final examination in biology. The other axis in the chart includes the major topics of course content. Such a chart of specifications should be prepared with great care before any item writing is undertaken. Then the chart can be followed as closely as the workmen follow the architect's blueprints when constructing a fine building. The proportion of items desired to test each unit of subject matter in terms of each objective should be determined in advance. As each item is completed, a tally mark or identifying symbol may be entered in the appropriate cell on the chart. The item writing and tallying should continue until the predetermined number of items for each cell on the chart has been completed. When the test is finished, the tally marks can be replaced by the numbers assigned to the items in the test. At the rate of approximately 10 items per week, the average science teacher will probably require perhaps 10 to 12 weeks to write the items or test exercises for a double-period semester final examination. About 50 items per class period of testing time would perhaps be appropriate for secondary school, and perhaps 30 items per class period for the upper elementary grades may be reasonable.

More difficult to evaluate are attitude, interest, and appreciation. These can be appraised more effectively in relation to projects, out-of-class activities, and situations that arise spontaneously in class than they can by formal testing.

These qualities should be indirectly reflected in student achievement. For example, if a student is *genuinely* interested in biological science this interest should motivate him to work diligently at *all* aspects of the course. Too frequently a tangential interest in one phase which carries the student far afield to the total neglect of many of the more fundamental aspects is mistaken for a genuine interest in the subject. Unless the interest reflects itself in consistently good performance throughout the course, it may well be that this marginal interest arose previously or quite independently of the course. The course grade will probably be more meaningful if the evaluation of the student is based primarily upon achievement that can be measured. However, this does not mean simple memorization.

It is helpful to check test questions for their accuracy and effectiveness. If several teachers are teaching the same course, construction of the test could be a team effort with division of labor and critical review of each other's items, followed by revisions where these seem appropriate. The team effort could be extended to assembling the test, writing the directions, proof-

reading the mimeograph stencils or ditto masters, and checking the accuracy of the answer key by having each team member go through the test and attempt to answer each item.

Answer Sheets

For an objective test a separate answer sheet facilitates scoring of the papers. If commercial answer sheets are not available, very serviceable answer sheets can be mimeographed or drawn up by hand on ditto masters.

Selecting Material for Questions

One of the ways a science test may involve the student's understanding of concepts as well as knowledge of facts is through the use of reading passages from magazines, newspapers, or scientific journals. Preparation of such items for a test is discussed later in this book, see page 22.

File Cards

If machine scoring and item analysis services are available, plan to file the best items, together with analysis data, on cards for possible use again in a subsequent test. A set of 5 x 8-inch cards, with headings similar to those shown in Figure II would serve as a convenient means of filing the best items. These cards could be run off either on the ditto or the mimeograph machine. For a large operation which might involve filing thousands of items, a printed card might be justified. Before re-using an item, however, weak foils (wrong answers), selected by very few or no students, should be made to seem more attractive. Conversely, foils that were selected by too many students should be made less attractive.

The Key Item

One structure for the test to be built of is the "key" item. This involves a series of multiple-choice item stems, all of which use a common set of alternatives from which answers are to be selected. Its chief advantage over using the same number of separate multiple-choice items is the greater compactness and economy of words and space. Care must be taken, however, lest irrelevant clues to the answer be inadvertently introduced.

In setting up the key, two fairly self-evident precautions should be observed. First, the key categories must be sufficiently *discrete* and *mutually-exclusive* so that there is no overlap. This is illustrated in these examples:

EXAMPLE 1

KEY:
1. Problem
2. Inference
3. Hypothesis
4. Assumption
5. Conclusion

Item 1. What induces birds to migrate? (This item is easily categorized as a problem.)

FIGURE I. Two-axis chart of specifications for an objective biology test. The numbers in each cell designate the items in a hypothetical test.

Content ▼ Objectives ▶	Cognitive Domain: Fundamental Knowledge	Intellectual Abilities and Skills: Comprehension	Application	Analysis	Total
Methodology of science		1, 2, 3, 4	5, 39-42	15-20	15
Characteristics common to living organisms. Cellular structure and functioning. Molecular biology	6, 7-11, 12, 13, 14, 20-23	24, 25	28-33		20
Kinds of living things and their groupings	26, 27, 28				3
Nature of processes essential to the life of individual organisms, including food manufacture, circulation, excretion, coordination, and adjustment. Parasitism and disease in man	34, 37, 43, 46, 47, 53, 54, 55, 56, 57, 58, 63, 101	35, 45, 48-51, 76-80	52, 67, 59-62	38	31
Processes associated with the continuance of the species—reproduction and heredity	68, 69, 70, 71	72, 73, 74	91-100	81-90	27
Ecological interrelationships. The world biome. Conserving natural resources. Economic plants and animals	44, 108, 115	107, 109, 110, 111, 112, 113, 116			11
Biology of space travel			64, 65, 66		3
History of life on the earth. Theories that attempt to account for evolutionary change.	75, 102	103, 104, 105, 106			6
Total	37	31	31	31	116

FIGURE II. Card for Filing Test Items

(A 5 x 8 inch card is recommended to allow adequate space for entering the various items of information.)

TOPIC	OBJECTIVE

Key	REFERENCE	When Used	No.	Discrim.	Difficulty

Item 2. Failure of the food supply in the northern regions with the coming of autumn provides the birds with the stimulus to migrate. (This item is difficult to categorize because numbers 2, 3, 4, and 5 in the key are not altogether mutually exclusive.)

EXAMPLE 2

KEY: 1. Problem
2. Hypothesis
3. Demonstrable fact
4. Fact by definition
5. Principle

Item 2 becomes more manageable with a more carefully refined key (Example 2), because now there is no overlap among the categories. They are mutually exclusive. In this context, item 2 is a hypothesis.

A second precaution to observe in setting up the key is to make the forms *homogenous* and *grammatically consistent*. All the categories should relate to one central idea; do not mix names, dates, processes, and definition terms in one key. Moreover, do not include in the same key proper names, common nouns, verbs, and adjectives. A single key should contain terms in only one category. The items should bear a relationship to each other, but at the same time each must be discrete and distinguishable from the others with no overlap. All responses must be grammatically consistent to be potential answers for any question.

Keys need not necessarily be made up of single words. Many "thought" questions require more elaborate keys. The following examples illustrate several ways of writing keys and statements.

EXAMPLE for use with statements

For items (give numbers) select the most appropriate response from the following key.

On the basis of the evidence

1. the statement can be considered a fact.
2. the statement is a reasonable hypothesis.
3. the statement is not warranted.

[Items]

EXAMPLE for use with a reading selection

Items (give numbers) are concerned with the following reading selection. Read it carefully. Then consider the statement(s):

[Reading selection]

1. Some diseases can be spread by people who are not ill.

In the light of the selection:

1. the statement is warranted.
2. the statement is not warranted.
3. the statement cannot be judged.

When two or more excerpts or passages on the same subject are to be compared or contrasted the following kind of key might be used.

EXAMPLE for use with two or more reading selections

Items (give numbers) are based on the following three reading selections. Please read these selections carefully.

[Reading selections A, B, and C]

Mark these items according to the following key.

1. Those who deliberately cause erosion or destruction of the topsoil to occur should be restrained.

The statement reflects the viewpoint of:

KEY: 1. A only.
2. B only.
3. C only.
4. all of these.
5. none of these.

True-False Tests

The ordinary two-alternative true-false type of test is looked upon with disfavor in many quarters, largely, perhaps, because it is frequently used to test for tidbits of memorized unrelated minutia. The following variation of true-false testing procedure using a key, however, can elicit answers that must be formulated on the basis of very careful thinking.

EXAMPLE

Use the following key to answer items 1-7.

KEY: 1. The statement is true if the condition applies.
2. The statement is true regardless of the condition.
3. The statement is false if the condition applies.
4. The statement is false regardless of the condition.
5. Impossible to determine without more data.

Statement	Condition
1. The force of attraction or repulsion between two charges is inversely proportional to the square of their distances	if the charges are quantitatively the same. (2)
2. The direction of the flow of electrons through a conductor is irreversible even	if the charges on the terminals are reversed. (3)
3. Two charged particles repel each other	if the particles approach closely enough. (5)
4. A negatively charged particle repels a positively charged particle	if the negative particle has the larger charge. (4)
5. Two charged objects repel each other	if both attract similarly charged objects. (1)
6. At constant volume the absolute temperature and pressure of a gas are directly proportional	if the temperature is expressed in degrees Centigrade. (3)
7. At constant temperature the absolute pressure and volume of a gas are inversely proportional	if the pressure is expressed in millimeters of mercury and the volume is expressed in cubic centimeters. (2)

TESTS FOR OBJECTIVES

I. RECOGNIZING AND APPRAISING ASSUMPTIONS

The ability to identify assumptions, to distinguish facts from assumptions, and to determine what effect certain assumptions may have on the validity of a conclusion should receive consideration in a science course. To test whether students have acquired these abilities, the "key" item serves well for use in the appraisal of the conclusion.

EXAMPLE

Items (give numbers) involve interpretation of data concerning the effects of grazing on the soil. Read the following passage carefully.

In an area which had been moderately grazed, a steel tube 18 inches in diameter and 12 inches long was driven 9 inches into the ground. The part of the tube that was above ground, 3 inches deep and 18 inches in diameter, was filled with water. It took eleven minutes for all the water to soak into the ground. Another tube of the same size was similarly placed in a nearby ungrazed area. In this location the water soaked into the ground in one and one-half minutes.

Conclusion: Water soaks into ungrazed land faster than it soaks into grazed land.

Certain assumptions have a bearing on the above conclusion. Evaluate the assumptions given in the test items according to the following key.

KEY: 1. An assumption which must be made if the conclusion is to be valid.
2. An assumption which, if made, would tend to invalidate the conclusion.
3. An assumption which is unrelated to the validity of the conclusion.
4. Not an assumption but a mere restatement of data.
5. Not an assumption but another conclusion.

1. The composition of the soil in the two areas was approximately the same. (1)

The following example also shows how to deal with assumptions in a testing situation.

EXAMPLE

Items 1-7 are concerned with the following situation.

One of the methods formerly used by geologists to determine the age of the oceans was a calculation based on the amount of salt (NaCl) in the ocean, and the amount added to ocean waters each year by the rivers that empty into the ocean. If this method of age determination is used, certain assumptions must be made.

Items 1-7 consist of a number of assumptions. Indicate by use of the key number that the assumption in the item is

KEY: 1. **necessary** for the determination of the age of the oceans and is probably true.
2. **necessary** for the determination but is probably false.
3. **not necessary** for the determination but is probably true.
4. **not necessary** for the determination and is probably false.

1. The salt concentration of the oceans is gradually increasing. (3)
2. Oceans have been on the earth since our planet was formed. (4)
3. Ever since its origin, the earth has revolved around the sun. (3)
4. The oceans now contain all the salt that has ever been added to them. (2)
5. The salts which rivers have carried to the oceans have all occurred in mineral form in rocks before they were dissolved by the river water. (3)
6. The proportion of the lithosphere existing above the ocean waters has been constant through the geologic ages. (4)
7. The continental masses have existed in essentially their present outline since the formation of the earth. (4)

A Situation at the Junior High Level

The following example from *A Test of Problem Solving* illustrates a less sophisticated level than the preceding—one that would be suitable for junior high school.

EXAMPLE

Items 1 and 2 refer to the following paragraph:

Differences in size, shape, composition, and characteristics of living things can be observed all around us. Even two peas in a pod vary in color, shape, and size. Plants vary from the giant trees to the tiny bacteria which can be seen only with the aid of a microscope. Animals vary from the huge whales to the one-celled organisms known as protozoa.

Conclusion: No two living things are exactly alike in all respects.

Items 1 and 2 are statements which might be made in a discussion of this paragraph and its conclusion.

Which of the following judgments would you make about the statements in Items 1 and 2?

KEY: 1. This statement tends to support the conclusion.
2. This statement tends to deny the conclusion.
3. This statement neither supports nor denies the conclusion.

1. People are often identified by their fingerprints. (1)
2. Sometimes a single fertilized egg splits and produces identical twins. (2)

A Situation at the Elementary School Level

Pupils in grade three, after having studied rules of safety, are given an opportunity to apply what they have learned in the following test situation:

EXAMPLE

Jack crosses the street at the corner only when the traffic light facing him has turned green. David looks both ways, and if he sees no cars he crosses the street whether the traffic light is green or red. Arthur dashes across the street in the middle of the block. Safety rules are best observed by

* A. Jack B. David C. Arthur

Evaluating Hypotheses

At this point we should make certain that the difference between a problem and a hypothesis is clearly recognized. A problem usually involves a perplexing question or situation that requires more than a "yes" or "no" answer or explanation. For example: *What is it that induces birds to migrate?* constitutes a problem. A prediction or proposed solution which provides a basis for further investigation constitutes a hypothesis. The hypothesis is usually cast in the form of a positive statement which can be either accepted or rejected when the evidence concerning it is examined and analyzed. In this respect the hypothesis is a "yes" or

"no" proposition—one that is supported or refuted by the evidence. Hypotheses are therefore tentative. Those that are refuted by the evidence are dropped for the time being from further consideration while those that are supported by the evidence become tentative conclusions subject to further testing.

Persons who possess a degree of statistical sophistication will point out that hypotheses are best expressed as null hypotheses which, when tested by certain statistical procedures, are accepted or rejected at stated levels of confidence, rather than being proved true or false. Anyone who wishes to pursue this point further can refer to a textbook in statistics.[1] For our purposes it is sufficient to be aware that evidence *tends to support* or *tends to refute* a given hypothesis—does not prove its absolute truth or falsity.

Being able to evaluate hypotheses is an important ability in science, and science courses aim to develop this objective. Questions can be devised that will allow the student to judge or evaluate hypotheses in relation to a problem, or make up his own hypothesis, or decide whether there is enough evidence to make a hypothesis. A good technique is to present a paragraph setting the context, followed by a statement or series of statements, and one or more hypotheses.

What is it that induces birds to migrate? Having posed the problem, we are ready to evaluate hypotheses in the light of data that bear on these hypotheses. One hypothesis and several items follow.

HYPOTHESIS: **Young birds learn to migrate south in the autumn by accompanying their parents.**

Each of the following items states observational or experimental facts. For items 1 and 2 mark space

a. if the fact or facts tend to support the hypothesis
b. if the fact or facts tend to refute the hypothesis
c. if the fact or facts are irrelevant to the hypothesis

1. In making the trip from northern Canada to the Argentine, adult golden plovers fly in groups due south from Labrador, above the Atlantic Ocean, 1,000 miles from the eastern coast of North America, thence over Brazil to the Argentine. The young golden plovers fly in groups by way of the Mississippi Valley and over the Gulf of Mexico, thence over Bolivia and Peru to the Argentine. On their return trip the old and young alike take the Mississippi Valley route. (b)

2. Cowbirds build no nests of their own but lay their eggs in the nests of some 30 species of other birds; here the eggs are incubated, hatched, and the young reared by these foster parents. When the first migratory trip is completed the young cowbirds are found in the South neither among their true parents nor scattered among the 30 different destinations of their foster parents, but instead in one locality occupied predominantly by their own kind. (b)

3. In the light of the facts given in items 1 and 2, what is the status of the hypothesis at this point?
   1. It is established as probably true
   *2. It is refuted as probably false
   3. It remains as much unsettled as at the outset

II. UNDERSTANDING OF SCIENCE METHODOLOGY

Some Classic Experiments

To determine whether students understand science methodology, it may be desirable to have them go through an analysis of a few relatively simple classical experiments. The following situations will serve as examples.

The work of Francesco Redi on spontaneous generation performed in the seventeenth century, stands out for its simplicity, clarity, and clean-cut design. This was one of the very earliest instances of a *control* being used in biological science to establish the validity of experimental results. Using a control as a standard for comparison is one of the most fundamental techniques in scientific experimentation.

The following set of five items requires careful analysis of the experiment that is described. Not many such items would be included in one test, but inclusion of a few will test the level of the students' ability to think clearly.

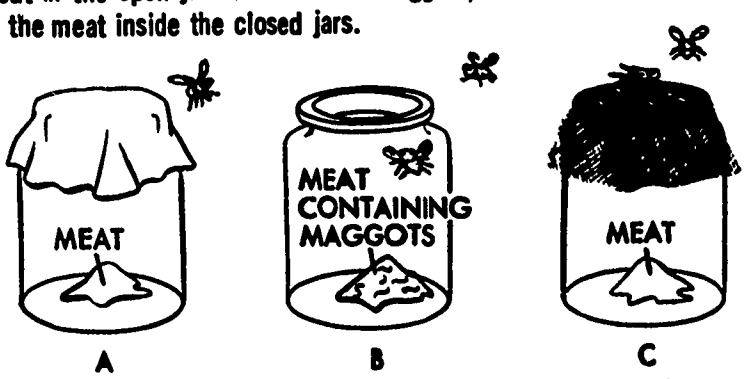
EXAMPLE

Items 1-5 are based upon the following situation.

PROBLEM: How do the simpler living organisms originate?

HYPOTHESIS: **Flies may be produced by spontaneous generation from dead organic substances.**

Francesco Redi, an Italian physician and scientist, put pieces of fresh meat inside one set of jars which he immediately sealed with parchment (illustrated by Jar A in the figure). inside another set of jars (illustrated by Jar B) he also put some pieces of fresh meat, but left these jars open. Later he observed flies entering and leaving the open jars at will. No flies could enter the closed jars. Some days later the meat in the open jars teemed with maggots, but no maggots developed in the meat inside the closed jars.



As a refinement in procedure, he repeated the experiment but instead of covering one set of jars with parchment he now closed them with fine-meshed gauze (illustrated by Jar C). Flies, possibly or perhaps

[1] These books may be helpful:

Mosteller, F.; Rourke, R. E. K.; and Thomas, G. B., Jr. *Probability and Statistics.* Paperback. Addison-Wesley Publishing Company, Reading, Massachusetts. 1961.

Youden, W. J. *Experimentation and Measurement.* VISTAS OF SCIENCE® paperback. National Science Teachers Association, Washington, D.C. 1962.

attracted by the odor of the flesh inside the jars, frequently alighted on the cloth, occasionally depositing eggs on it. These eggs soon hatched into maggots on top of the gauze, but no maggots developed in the meat inside the jars.

For items 1-4 mark space

1. if the item is true according to the data and tends to support the hypothesis;
2. if the item is true according to the data but tends to refute the hypothesis;
3. if the item is irrelevant to the hypothesis, regardless of its truth or falsity according to the data;
4. if the item is false according to the data, but if true, would tend to support the hypothesis;
5. if the item is false according to the data, but if true, would tend to refute the hypothesis.

SUGGESTION: It will be easier if you decide first whether each statement is true or false according to the data.

1. Judging on the basis of what happened in all three jars, no maggots were to be seen in the meat in Jar A because they suffocated in this tightly closed jar. (4)
2. Jar B, which was the only jar that the flies could enter, was also the only jar in which maggots appeared in the meat. (2)
3. The maggots which appeared on top of the gauze of Jar C appeared there only because tiny particles of decaying meat that were carried upward through the circulating air turned into maggots. (4)
4. The reason the maggots were not found in Jars A or C is the same. (2)
5. On the basis of Redi's data alone, what is the status of the hypothesis?
   1. It is established as true beyond doubt.
   2. It is probably true—the evidence tends to support it.
   3. It remains as much unsettled as at the outset.
   *4. It is probably false—the evidence tends to refute it.
   5. It is definitely false without any doubt whatsoever.

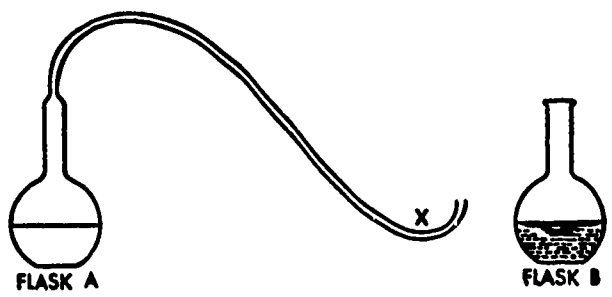
In the second half of the nineteenth century Pasteur shed further light on the problem of spontaneous generation when he, too, employed the technique of controlled experimentation in solving a troublesome phase of this problem. His method also was characterized by simplicity, clarity, and clean-cut experimental design. A classic in the field of science, Pasteur's experiment can very well serve as the basis for an analytical testing situation.

Because of the length of both the preceding and the following set, one ought not to include both the Redi situation and the Pasteur experiment in the same test.

EXAMPLE

Items 1-8 are based upon the following passage which describes some of Pasteur's later experimental work on spontaneous generation.

After some preliminary experimentation Pasteur developed a refinement in his method which was designed to meet various objections raised by his critics. This time he boiled the sugared yeast water in a swan-neck flask like flask A in the illustration.



Pasteur's swan-neck flask (A) and open-top flask (B)

The sugared yeast water in both flasks was boiled, then allowed to cool and stand for several weeks. Air could enter flask A via the long bent tube. The contents of flask B were exposed to the air through its open top. A dense clouding soon appeared in flask B, but none developed in flask A.

Pasteur allowed the contents of flask A to cool very slowly. The end of the stem was not sealed. Any air entering the swan-neck was believed to have lost its load of dust, microorganisms, and other suspended debris when it began ascending the long uphill part of the swan-neck. In all likelihood only air free of microorganisms came into contact with the sugared yeast water inside the swan-neck flask. No clouding appeared in the flask even after months of standing.

Pasteur believed that any microorganisms borne by the air entering the swan-neck tube may have lodged in the lowermost curve near the entrance of the tube. To find out if this were indeed true, he tilted the flask in such a way as to allow some of the fluid within the flask to run down into the portion of the curve at "X" in the sketch. When this fluid was subsequently drained back into the flask and the flask had been incubated for a week longer, clouding of the formerly clear liquid in the flask gradually occurred.

For items 1-3 select from the key the most appropriate category.

KEY:
1. Experimental observation
2. Generalization
3. Analogy
4. Deduction
5. None of the above

1. The liquid in flask A remained clear even after months of standing. (1)
2. Whenever sugared yeast water is kept sterile it remains clear, but whenever microorganism-laden air comes into contact with sugared yeast water, the liquid becomes clouded. (2)
3. Since the liquid in flask A remained clear, it contained no growth of microorganisms. (4)
4. Pasteur allowed the contents of flask A to cool very slowly in order to
   1. avoid the danger of an explosion inside the flask.
   *2. prevent an inrush of air and microorganisms.
   3. avoid shock to the microorganisms by sudden temperature change.
   4. prevent the flask from cracking and spilling its contents.
   5. give himself time to make observations while the contents of the flask were cooling.
5. By this experiment Pasteur
   1. proved that spontaneous generation has never occurred on the earth.
   2. proved that microorganisms arise by spontaneous generation in a culture medium if the experimenter is careless, but they do not arise if he is careful.
   3. demonstrated that no culture medium, if it has been boiled, will become contaminated with microorganisms even if germ-free air comes into contact with the fluid inside the flask.
   *4. obtained evidence that sterile sugared yeast water in contact with germ-free air does not engender microorganisms.

To illustrate the point of view that scientific methodology is similar whether the experiment was performed many centuries ago or quite recently, we can compare the following analysis of an experimental situation with the two preceding examples. All three are concerned with the solution of a problem, all involve hypotheses, assumptions, and data gained from em-

pirical observation, and all three employ scientific reasoning, particularly of the deductive type.

EXAMPLE

Items 1-6 are concerned with the following situation:

The following selection is adapted from an address by Thomas Midgley, Jr. on how the gas called "freon 21" ($CHCl_2F$) was developed for mechanical refrigerators in the early 1930's.

One morning the director remarked to me that the refrigeration industry needed a new refrigerant if it ever expected to get anywhere. I was skeptical that anything other than a mixture of substances would reduce existing hazards, but after discussing it with the chief engineer, two of my associates and I went into the library and started working.

The desired combination of properties was a boiling point between 0°C and —40°C, stability, nontoxicity, and nonflammability. International Critical Tables gave us a partial summary of the volatile organic compounds. The now-proved mistake that carbon tetrafluoride boiled at —15°C struck us in the face and started us thinking about fluorine.

Recognizing that the Critical Table list was very incomplete, I decided to bring into play the periodic table. Perhaps volatility could be related to it in some way, and it took but a moment to see that this was true. Volatile compounds of boron, silicon, phosphorus, arsenic, antimony, bismuth, selenium, tellurium, and iodine (italicized in the accompanying table) are all too unstable and toxic to consider. The inert gases (starred in the accompanying table) are too low in boiling point.

							He*
...	...	*B*	C	N	O	F	Ne*
...	...	...	*Si*	*P*	S	Cl	A*
...	...	...	...	*As*	*Ce*	Br	Kr*
...	...	...	...	*Sb*	*Te*	*I*	Xe*
...	...	...	...	*Bi*	...	...	...

Now look over the remaining elements. Every refrigerant used has been made from combinations of these elements. Flammability decreases from left to right. Toxicity (in general) decreases from the heavy elements at the bottom to the lighter elements at the top. These two considerations focus on fluorine. It was an exciting deduction. Seemingly no one previously had considered it possible that fluorine might be nontoxic in some of its compounds.

The heats of formation between the halogens (fluorine, chlorine, bromine, iodine) and carbon were checked. They increased from iodine to fluorine, thus indicating a high degree of stability for fluorine-carbon compounds. Next came methods of preparation. Carbon tetrafluoride (CF_4) seemed rather hard to make. And then how could dichloro-difluoro-methane (CCl_2F_2) boil at —20°C and carbon tetrafluoride at —15°C? It just didn't make sense. After investigation, we decided that carbon tetrafluoride boiled at about —136°C. A publication on the subject appeared soon afterwards, reporting that carbon tetrafluoride boils at —128°C.

We selected dichloro-monofluoro-methane ($CHCl_2F$, usually called "freon 21") as the starting point for experimentation. I called one of the chemical supply houses by telephone and ordered five 1-ounce bottles of antimony trifluoride (SbF_3). One was taken at random, and a few grains of dichloro-monofluoro-methane were prepared. A guinea pig was placed under a bell jar with it, and much to the surprise of the physician in charge, didn't suddenly gasp and die. It wasn't even irritated. Our predictions were fulfilled. We took another bottle, made a few more grams, and tried it again. This time the animal did what the physician expected. We repeated again but this time we smelled the material first, and obviously it was contaminated. The answer was **phosgene.** Then we examined the two remaining bottles of antimony trifluoride and found that they were both badly contaminated with a double salt containing water of crystallization which resulted in the phosgene.

Of five bottles marked "antimony trifluoride," one had held good material, and we had chosen it by accident for our first trial. Had we chosen any one of the other four, the animal would have died as expected by everyone else in the world except ourselves. I believe we would have then given up.

For items 1-5 select the best answer.

1. The statement, "The heats of formation between the halogens and carbon were checked. They increase from iodine to fluorine," represents
   1. a hypothesis.
   2. a problem.
   3. the specification of a problem.
   *4. data used in the solution of a problem.
   5. none of these.
2. The statement, "And then how could dichloro-difluoro-methane boil at —20°C and carbon tetrafluoride at —15°C?" represents
   *1. a problem.
   2. a test of a hypothesis.
   3. a hypothesis.
   4. a fact by observation.
   5. a conclusion from systematic observation.
3. The statement, "A guinea pig was placed under the bell jar with it and . . . didn't suddenly gasp and die," represents
   1. a hypothesis.
   *2. a test of a hypothesis.
   3. a problem.
   4. a conclusion based on observations.
   5. generalized data.
4. The primary objective of this investigation was to
   1. determine the boiling points of fluorine compounds.
   2. correct errors in the International Critical Tables.
   *3. find a nonpoisonous, nonflammable refrigerant.
   4. determine the toxicity of dichloro-monofluoro-methane.
   5. test for impurities in antimony trifluoride.
5. In deciding to concentrate upon compounds of fluorine, which one of the following was a basic assumption?
   *1. Trends discovered in part of the Periodic Table are likely to continue.
   2. All compounds of a toxic element were known to be nontoxic.
   3. Compounds with lower boiling points are more flammable than those with higher boiling points.
   4. Mixtures of compounds are always less toxic than single compounds.
   5. Compounds of carbon are nonflammable.

(There were 11 additional items in the above set.)

Formulating and Testing Hypotheses

It is often desirable to present a problem to the class and let the students try to formulate some hypotheses at the outset. Since many of their hypotheses will probably be very similar to each other, the entire output of the class can probably be reduced to four or five. These can then be subjected to testing—either by experimentation, or by reasoning based on experimentation that others have done, or by empirical observation, or by a combination of these three methods.

As an example, some oat seeds had been planted in a seed flat. Within a few days after the seeds had germinated, the seed flat was placed where light shone on it from one side only. Soon afterward the seedlings were all bending toward the light source. The problem was then posed to the students: "What causes a green plant, exposed to the light on only one side, to bend toward the light as it grows?" The students were asked

to formulate a hypothesis and to write it on a slip of paper. These papers were immediately collected. Following is a sample of a few of their hypotheses:

When young seedlings are exposed to a bright light source on only one side, they tend to bend toward the light source as they grow because

- ...the part nearest the light grows faster and thus seems to bend; the part of the plant away from the source of light does not grow as fast, thus it doesn't seem to bend at all.
- ...light is needed for photosynthesis, and the plant is responding to the stimulus which is the sun. The plant responds to the light of the sun.
- ...the cells toward the light are engaged in photosynthesis, while cells away from the light are engaged in storing the sugar produced, and using it for growth. This causes the dark cells to grow faster, which causes the plant to bend toward the light, or away from the stress.
- ...plants need light in order to make food. When there is no light they cannot make food. Therefore it would seem there would be some mechanism in the plant to make it turn to the light.
- ...the plant is attracted by the light. When the plant grows at night it bends where the sun has bathed the stem to cover it, making it look as though it grows toward the light.
- ...the side of the plant in the dark grows faster and therefore makes the plant bend toward the light.
- ...if the plant bends there must be some growth taking place and a plant doesn't grow during the day but at night so therefore the shaded side would grow causing the plant to bend.
- ...it is their source of food. The plant responds to the warmth and light and bends toward it.

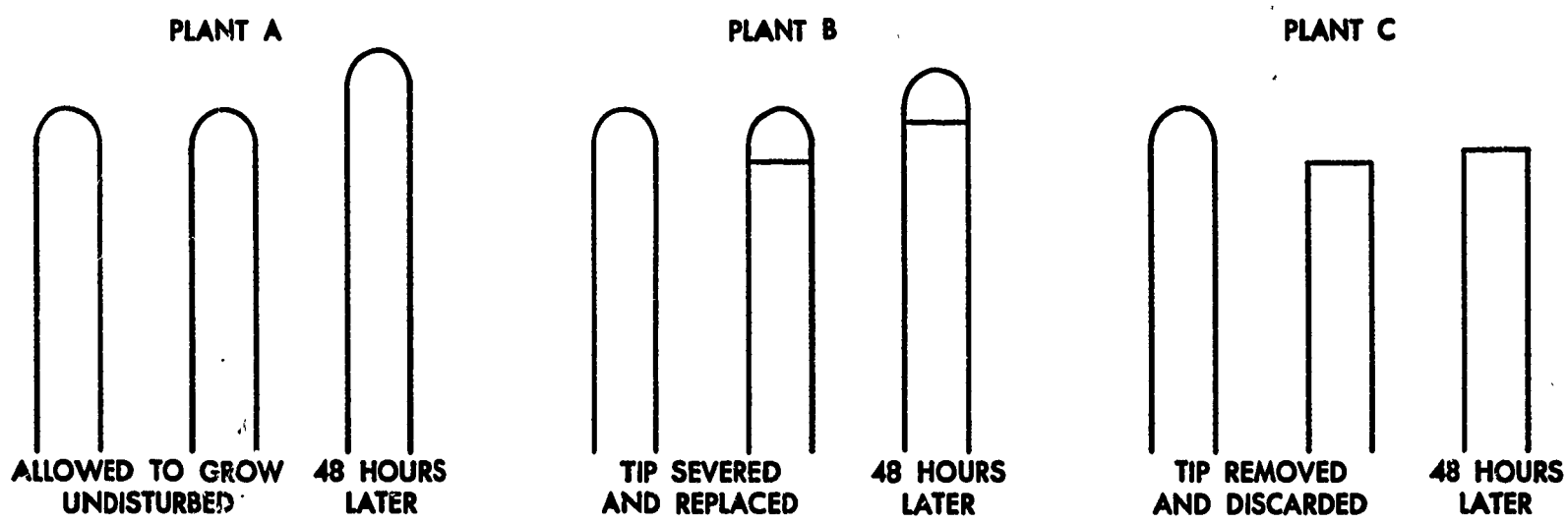
Since the remaining hypotheses submitted by the class were quite similar to these, it is not necessary to list them all.

The class then studied a series of hypotheses according to a procedure outlined in the laboratory manual. Later the students took an examination that included the following series of questions which closely paralled their laboratory study.

Items 1-15 are concerned with the following problem:
PROBLEM: **What causes plants to bend toward the light when they grow?**

Young oat seedlings were treated as indicated in the sketches accompanying each experiment. The seedlings in Experiments I-IV were grown in the dark for 48 hours.

EXPERIMENT I



1. What does the above experiment show?
   1. Plants need light in order to grow.
   2. Cutting off its tip encourages lateral growth in a plant.
   3. Chlorophyll is necessary for photosynthesis to occur.
   *4. Something present in the plant tip apparently stimulates the plant to grow in length.
   5. Cutting off the plant tip is a common horticultural practice which produces more satisfactory plant growth.
2. What would tend to preclude the possibility of a nerve connection between the plant tip and the rest of the stem?
   1. Because a plant has no brain it consequently has no nerves either.
   2. Possession of nerves would present a constant hazard in the life of a plant.
   3. Plants would have no use for nerves even if nerves were present.
   4. Nerves could not conduct substances in solution from one place to another.
   *5. The nerve connection between the tip and the rest of the stem would have been destroyed when the tip was severed.

to formulate a hypothesis and to write it on a slip of paper. These papers were immediately collected. Following is a sample of a few of their hypotheses:

When young seedlings are exposed to a bright light source on only one side, they tend to bend toward the light source as they grow because

- ... the part nearest the light grows faster and thus seems to bend; the part of the plant away from the source of light does not grow as fast, thus it doesn't seem to bend at all.
- ... light is needed for photosynthesis, and the plant is responding to the stimulus which is the sun. The plant responds to the light of the sun.
- ... the cells toward the light are engaged in photosynthesis, while cells away from the light are engaged in storing the sugar produced, and using it for growth. This causes the dark cells to grow faster, which causes the plant to bend toward the light, or away from the stress.
- ... plants need light in order to make food. When there is no light they cannot make food. Therefore it would seem there would be some mechanism in the plant to make it turn to the light.
- ... the plant is attracted by the light. When the plant grows at night it bends where the sun has bathed the stem to cover it, making it look as though it grows toward the light.
- ... the side of the plant in the dark grows faster and therefore makes the plant bend toward the light.
- ... if the plant bends there must be some growth taking place and a plant doesn't grow during the day but at night so therefore the shaded side would grow causing the plant to bend.
- ... it is their source of food. The plant responds to the warmth and light and bends toward it.

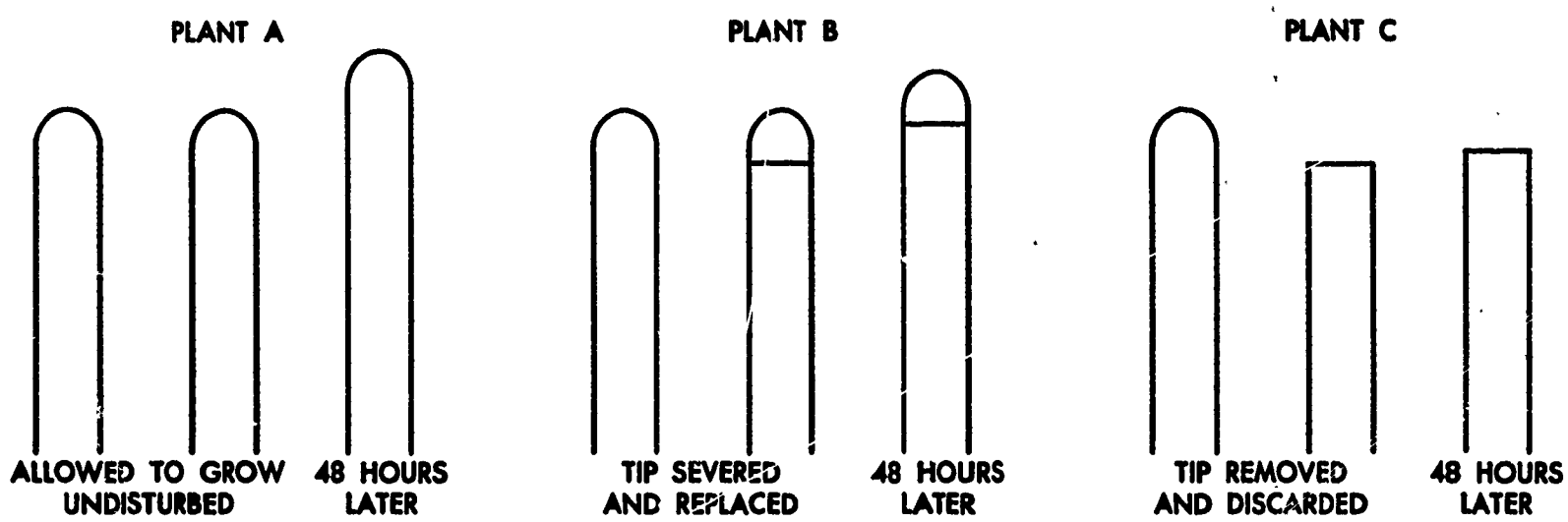
Since the remaining hypotheses submitted by the class were quite similar to these, it is not necessary to list them all.

The class then studied a series of hypotheses according to a procedure outlined in the laboratory manual. Later the students took an examination that included the following series of questions which closely paralled their laboratory study.

Items 1-15 are concerned with the following problem:
PROBLEM: **What causes plants to bend toward the light when they grow?**

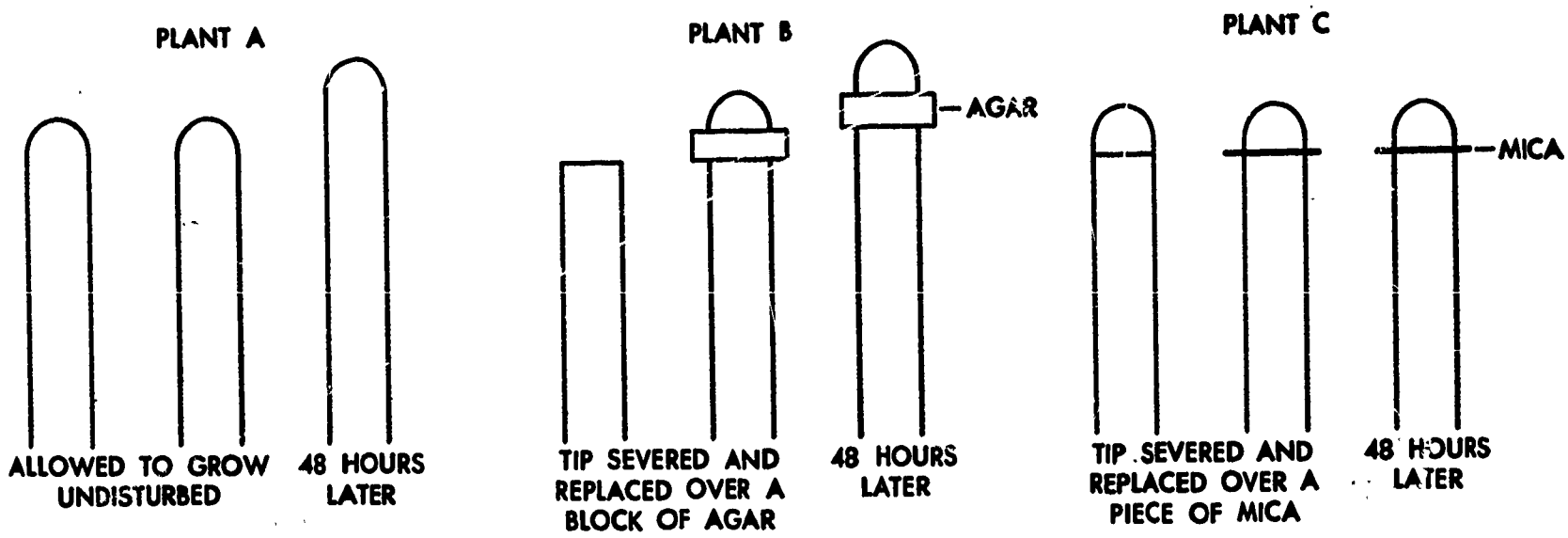
Young oat seedlings were treated as indicated in the sketches accompanying each experiment. The seedlings in Experiments I-IV were grown in the dark for 48 hours.

EXPERIMENT I



1. What does the above experiment show?
   1. Plants need light in order to grow.
   2. Cutting off its tip encourages lateral growth in a plant.
   3. Chlorophyll is necessary for photosynthesis to occur.
   *4. Something present in the plant tip apparently stimulates the plant to grow in length.
   5. Cutting off the plant tip is a common horticultural practice which produces more satisfactory plant growth.
2. What would tend to preclude the possibility of a nerve connection between the plant tip and the rest of the stem?
   1. Because a plant has no brain it consequently has no nerves either.
   2. Possession of nerves would present a constant hazard in the life of a plant.
   3. Plants would have no use for nerves even if nerves were present.
   4. Nerves could not conduct substances in solution from one place to another.
   *5. The nerve connection between the tip and the rest of the stem would have been destroyed when the tip was severed.

EXPERIMENT II



3. At this stage, four of the following are possible explanations of what has happened in Experiment II. Which one could **not** be accepted as an explanation at this point?

   1. Agar may have a stimulating effect on plant stem growth.
   2. Mica is impermeable to certain substances present in plants.
   3. Regeneration of some sort of connection may occur through agar but not through mica.
   4. Something from the stem tip appears to diffuse down through the agar and stimulate stem growth.
   *5. A plant will die if its stem tip is cut off.

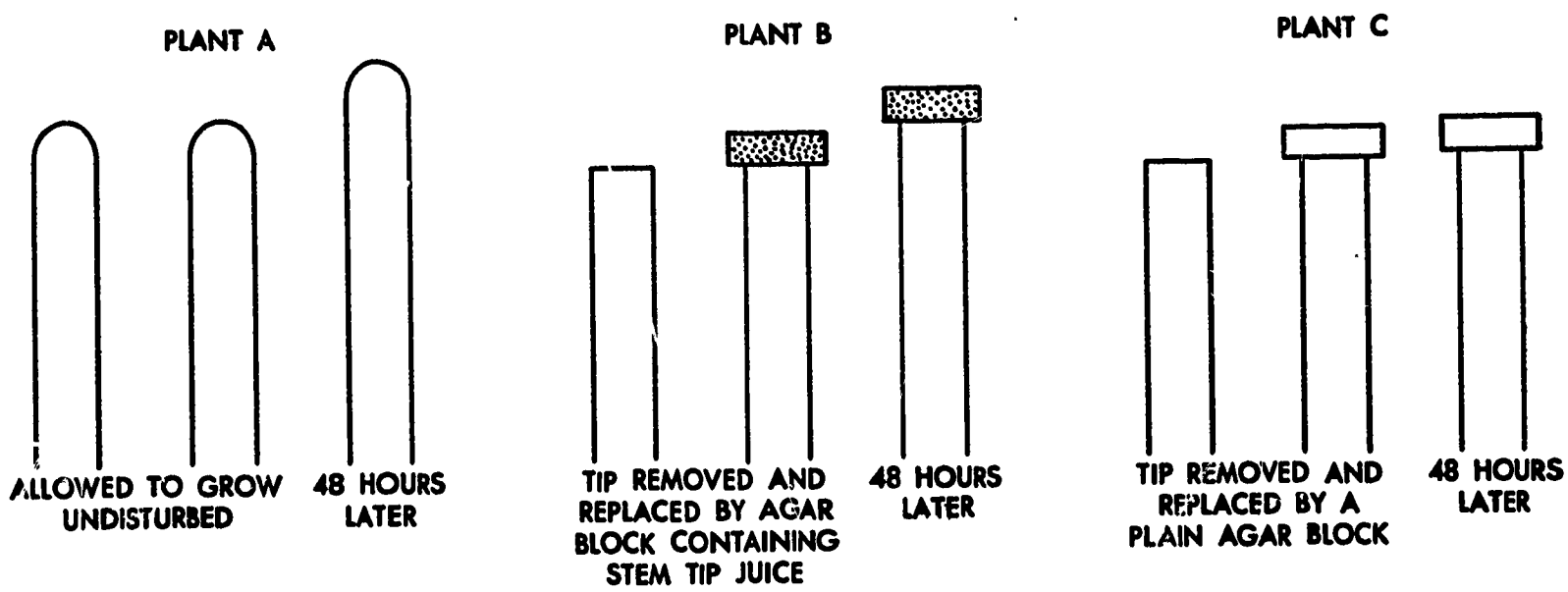
4. Which plants indicated that a diffusible substance may be present in the stem tip?

   1. Plants A alone
   2. Plants B alone
   3. Plants C alone
   *4. All three sets of plants
   5. None of the three sets of plants

5. Which plants indicated most clearly that if there is a diffusible substance in the tip it is ineffective unless it can diffuse downward?

   1. Plants A alone
   2. Plants B alone
   *3. Plants C alone
   4. All three sets of plants
   5. None of the three sets of plants

EXPERIMENT III



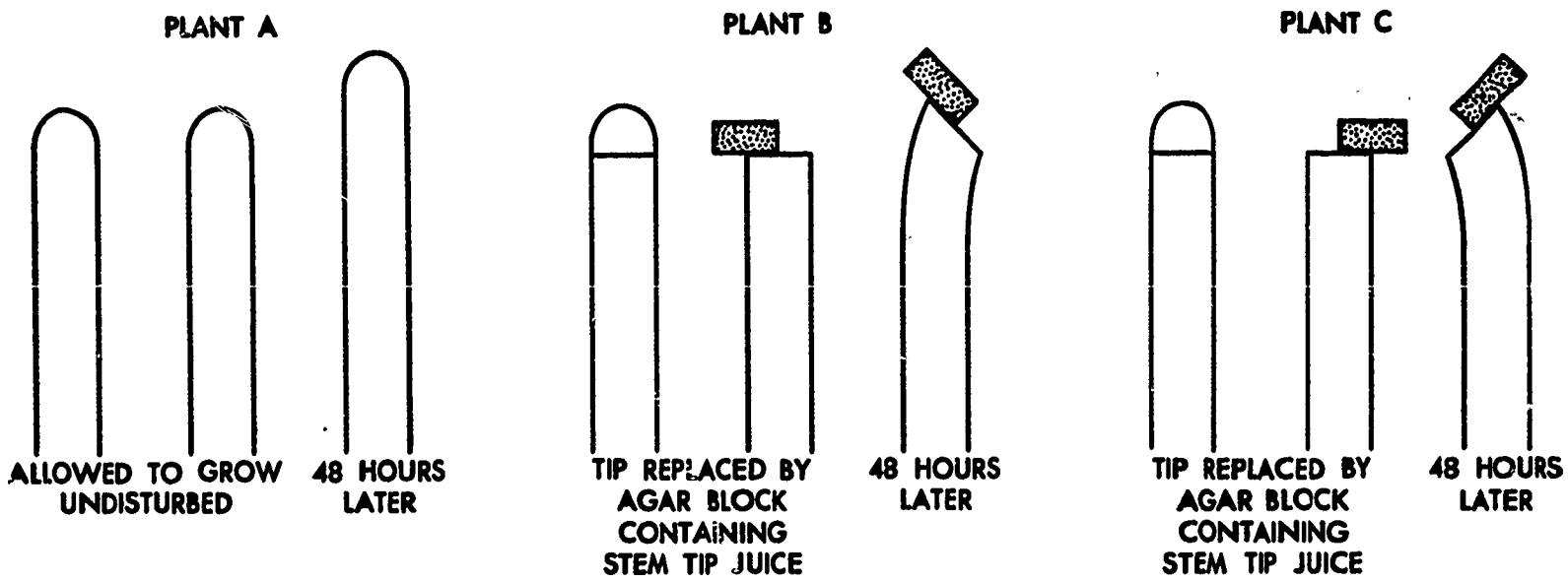
6. Which of the following is warranted on the basis of Experiment III?

   1. Agar has a stimulating effect on plant growth.
   2. A plant will die if its stem tip is cut off.
   *3. A substance in the juice from the stem tip is capable of diffusing through the agar.
   4. A plant will elongate whether or not it is exposed to light.
   5. The stem tip has little or no effect upon plant growth.

7. Experiment III **eliminates** which one of the following as a possible explanation of what happened in Experiment II?

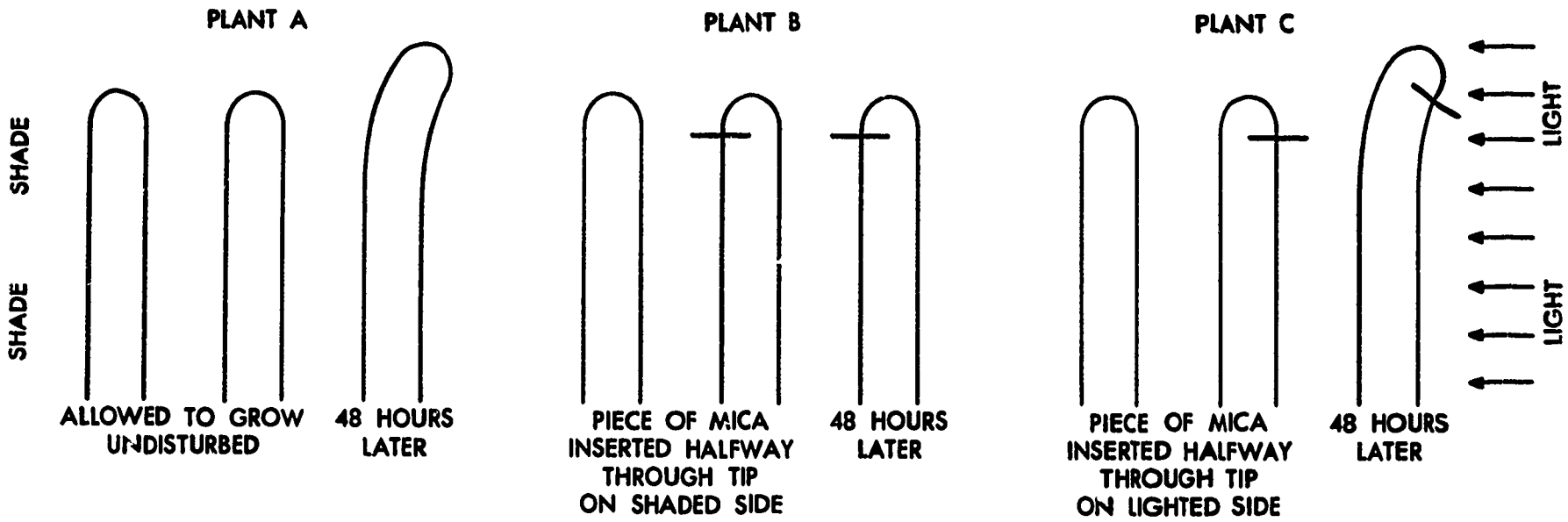
   *1. Agar has a stimulating effect on plant growth.
   2. Mica has a toxic effect on plants.
   3. A plant must have light in order to increase in length.
   4. Removal of the stem tip brings photosynthesis to a halt.
   5. None of the above is eliminated up to this point.

EXPERIMENT IV



8. Which plants indicated most clearly that the diffusible substance very probably diffused vertically downward but not laterally when the plants are grown in darkness for 48 hours?
   1. Plants A alone
   2. Plants B alone
   3. Plants C alone
   *4. Plants B and C
   5. All three sets of plants
9. The bending occurred in plants B and C but not in plant A because:
   1. Plants in A lacked a growth hormone.
   2. Agar is a growth stimulant.
   3. Elongation of cells was retarded by growth hormone.
   4. The cells grew faster in B and C plants than in A plants.
   *5. Greater elongation of cells took place on the side in which downward diffusion of hormone occurred.
10. On the basis of the evidence in this experiment, why did the stems of plants B3 and C3 bend as they elongated?
   1. They were attracted to the light.
   2. Stem-tip substance probably diffused laterally causing more rapid growth on the side toward which the stem bent.
   3. The agar had a stimulating effect on the cells directly underneath the agar block in each case.
   *4. Stem-tip substance probably diffused vertically downward causing greater cell elongation on that side.
   5. The pressure of the agar block stimulated the cells just underneath it to grow faster.

EXPERIMENT V



11. What new experimental condition was introduced in Experiment V?
   1. A piece of mica was used.
   2. Stem-tip fluid was present in the experimental plant(s).
   3. The controls and experimental plants were reversed.
   4. A control was lacking.
   *5. Light.
12. Which of the following might be an explanation of the apparent effect of the light in Experiment V?
   1. It stimulated the plants to grow faster on the lighted side.
   2. It was necessary for maximum plant stem elongation.
   3. It made unnecessary the removal of the stem tip.
   *4. It may have caused a dispersal or inhibition in effectiveness of stem tip substance.
   5. It produced no noticeable effect.
13. Why did plant B fail to curve?
   1. The mica had a toxic effect.
   2. Growth hormone could not reach the stem tip from below due to the obstruction of the mica.
   3. Plant B was unable to synthesize any growth hormone.
   *4. Growth hormone could not diffuse downward past the mica on the shady side.
   5. There is no logical explanation—it should have curved to the left.

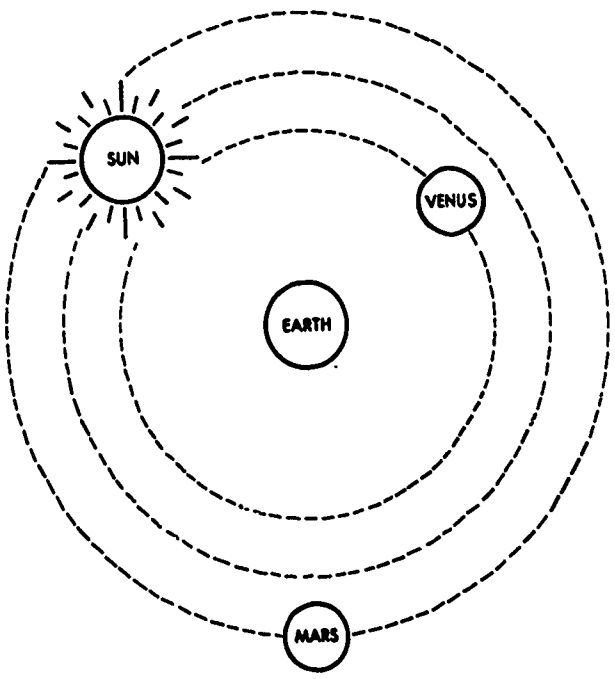
14. If a Midwest farmer were to measure the lengths of 10 cornstalks in June at 6 AM and at 6 PM on five successive days and compute the average daytime increase in length as well as the average nighttime increase in length, which of the following would he be likely to find?
    1. The average daytime increase in length was greater than the average nighttime increase in length.
    2. The average daytime increase and the average nighttime increase were about the same.
    *3. The average nighttime increase in length was greater than the average daytime increase in length.
    4. There was no increase in length during the night.
    5. Some stalks grew faster during the daytime than at night, while others showed greater elongation during the nighttime than during the day.

15. Which of the following constitutes the best explanation as to what causes a green plant exposed to the light on only one side to bend toward the light as it grows?
    1. Green plants need light to carry on photosynthesis.
    2. Green plants are phototropic.
    3. Light stimulates plant cells to grow faster.
    *4. More growth hormone accumulates on the shady side, stimulating greater cell elongation on that side.
    5. The light attracts the plant by stimulating more rapid growth in that part of the plant on which the light falls.

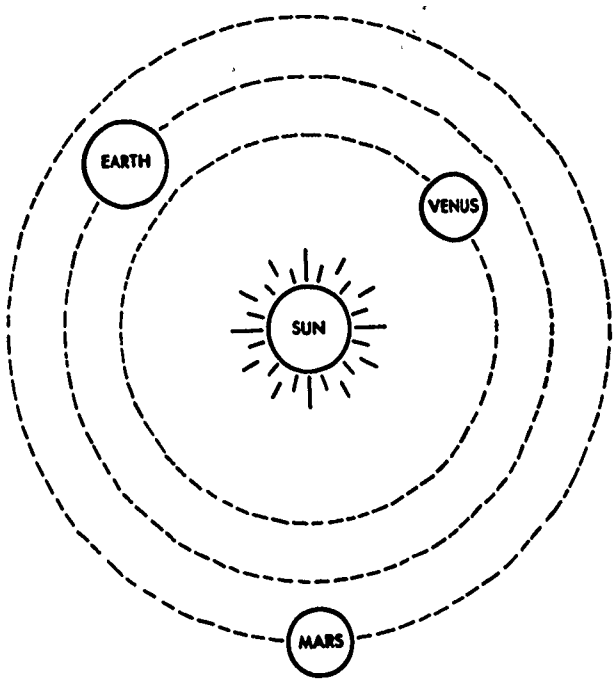
Comparing and Contrasting Two Theories

Analysis of scientific theories occurs quite frequently in the classroom. On occasion it may be desirable to compare and contrast two theories that attempt to explain the same set of phenomena. The typical testing procedure which follows such an activity is to ask the students to write several paragraphs comparing and contrasting the two theories. The following items show how objective testing techniques can be employed to accomplish the same purpose. Study the following diagram carefully.

PTOLEMAIC THEORY



COPERNICAN THEORY



Items 1-7 are based upon the following situation. Please read the description carefully.

In the sixteenth century, scholars debated the relative merits of two theories of the solar system, which are diagramed in the simplified form in the figure. The PTOLEMAIC theory assumed that the sun and planets revolved in concentric, circular orbits around a centrally placed earth. The COPERNICAN theory assumed that the earth and planets revolved in concentric, circular orbits around a centrally placed sun. BOTH theories postulated a region of fixed stars just beyond the outermost orbit.

Important observations bearing on this controversy were made possible by the invention of the telescope. Items 1-7 are examples of these observations. For each item, select the most appropriate category from the following key.

KEY: 1. Observation is favorable to the COPERNICAN theory and unfavorable to the PTOLEMAIC theory.
2. Observation is favorable to the PTOLEMAIC theory and unfavorable to the COPERNICAN theory.
3. Observation is predictable on the basis of either theory and is therefore favorable to BOTH.
4. Observation is contrary to the expectations of either theory and is therefore unfavorable to BOTH.
5. Observation does not bear on the validity of either theory as expressed in the description.

1. Mars was found to have two satellites revolving around it in the same direction. (5)
2. As viewed from the earth both Venus and Mars sometimes pass behind the sun. (1)
3. Venus, but not Mars, sometimes passes between the earth and the sun so as to eclipse a part of the sun. (3)
4. Venus goes through phases (full, half, crescent, etc.). Moreover, there is a size-variation associated with phase-variation. The apparent diameter of the planet, as seen from the earth, is greatest in crescent and least in full phase. (1)

5. Mars also seems to vary greatly in diameter as seen from the earth. (1)
6. Careful measurements disclosed that even the sun shows an apparent variation in size from season to season. (5)
7. Careful measurements, however, could not detect any variation in sizes, positions, or distances between any of the fixed stars as seen from the earth at different seasons of the year. (2)

Analyzing Data in Relation to a Problem Situation

Preliminary to the solution of most problems it is essential to assess what is given in the data at the outset, what information is not given but can be inferred from that which is given, and what is given but irrelevant.

After examining the data it is also important to be able to determine what conclusions are supported by the data and what conclusions are contradicted by the data. Examples which consist of problem situations and an accompanying key and items that embody this type of activity are presented in the following exercises.

Items 1-7 are concerned with the following situation.

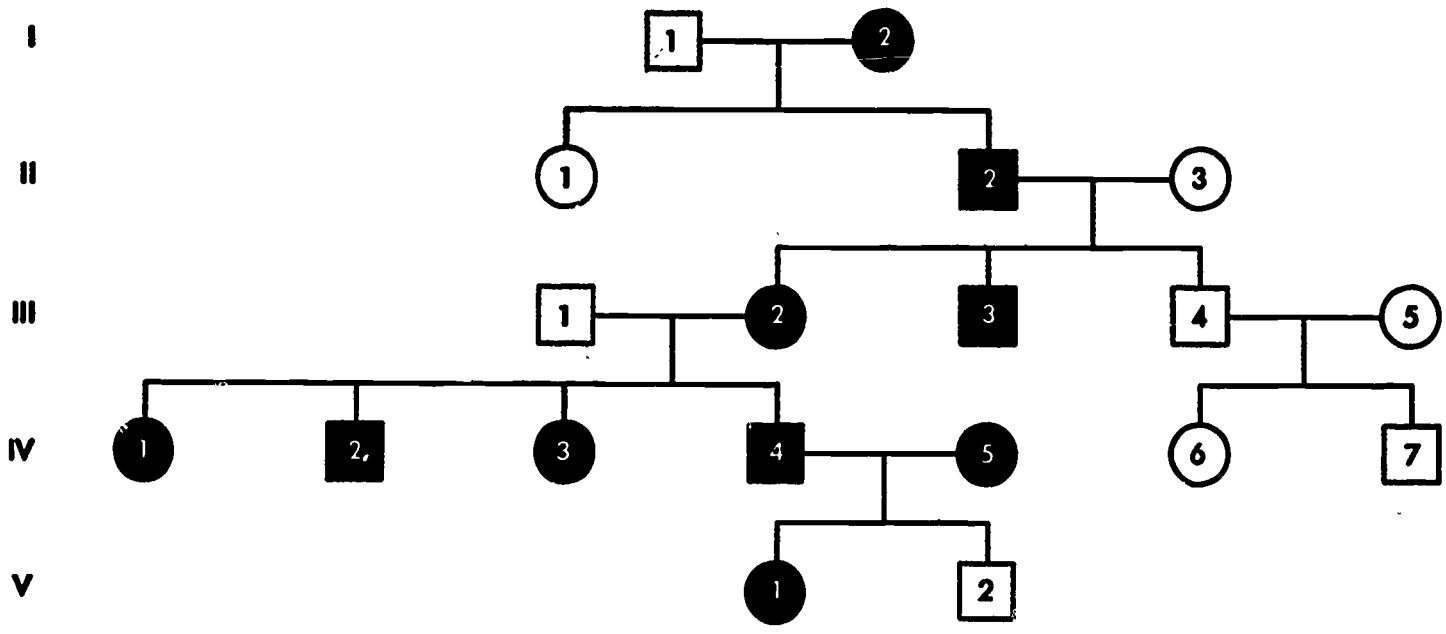
Above is a human pedigree chart showing the inheritance of premature whitening of the hair through five generations. Black squares represent affected males, and black circles represent affected females.

PROBLEM: **What is the inheritance pattern of the trait premature whitening of the hair?**

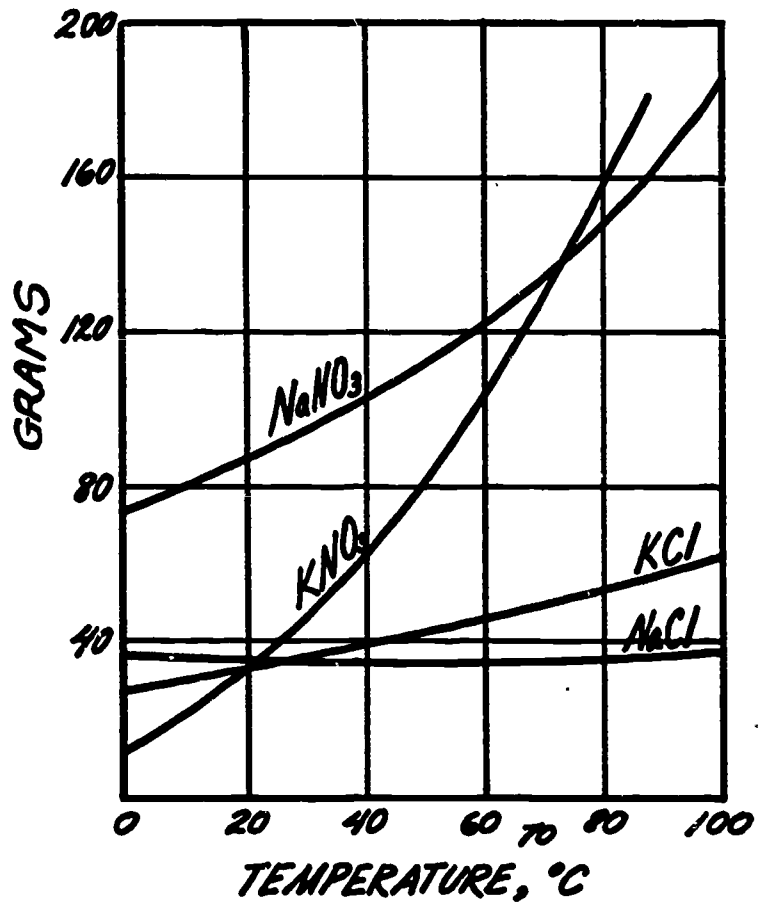
For items 1-7 select the most appropriate response from the key.

KEY: 1. Given in the data but irrelevant to the solution of the problem
2. Can be inferred from the above data
3. A conclusion that is contradicted by the above data
4. Given in the data and necessary to the solution of the problem
5. None of the above

1. Black squares and circles in the pedigree chart represent individuals who show premature whitening of the hair. (4)
2. The gene for premature whitening of the hair is sex-linked. (3)
3. The gene for premature whitening of the hair is recessive to the gene for normal hair pigmentation. (3)
4. The normal individuals represented in the pedigree chart are homozygous for normal hair pigmentation. (2)
5. The gene for premature whitening of the hair is allelic to the gene for normal hair pigmentation in the above pedigree. (2)
6. Individual III-3 is heterozygous for the trait in question. (2)
7. The cause of premature whitening of the hair may have some relationship to the cause of albinism. (5)



Refer to the diagram below. Mark each of the statements according to the key.



PROBLEM: **What are the relationships between the solubility and temperature for various compounds shown in the accompanying graph?**

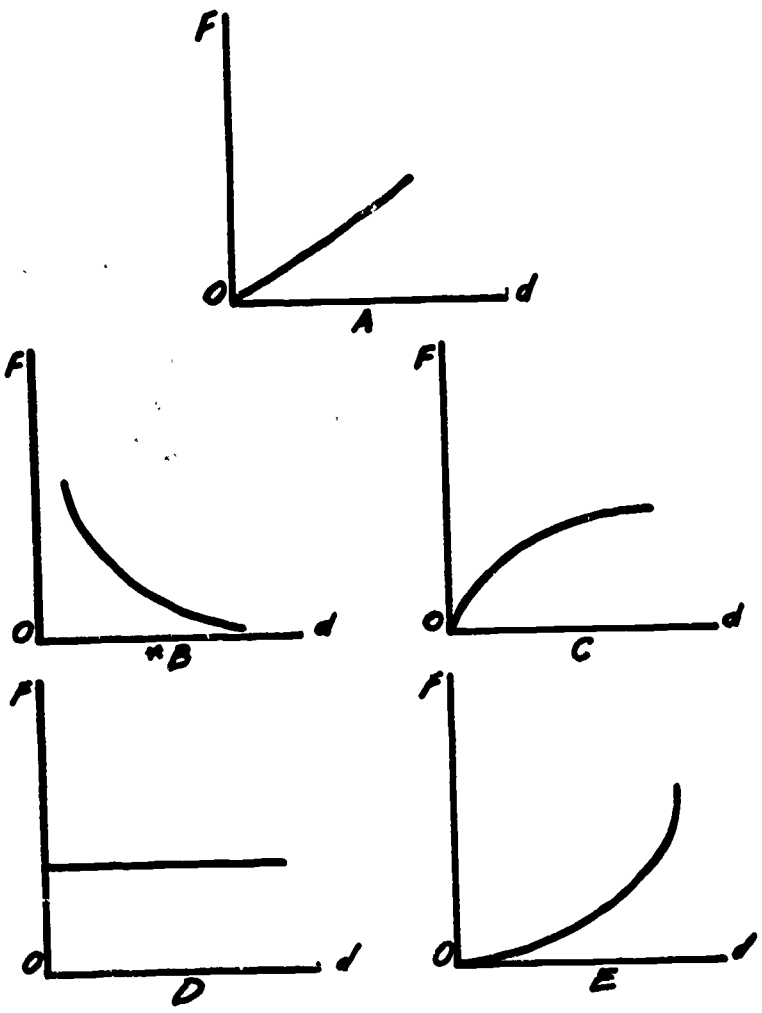
KEY: 1. True, the statement is supported by the data in the diagram.
2. False, the statement is disproved by the data in the diagram.
3. The data in the diagram neither prove nor disprove the statement.

1. Sodium nitrate and potassium nitrate have the same solubility at a temperature somewhat below 70°C. (1)
2. Potassium chloride is more soluble than sodium chloride at all temperatures. (2)
3. If a solution containing sodium, potassium, nitrate, and chloride ions is saturated at 100°C and is then cooled to 60°C, some sodium chloride probably will crystallize out. (1)
4. If a solution is saturated with both sodium nitrate and potassium chloride at 30°C and is then cooled to 10°C, some potassium nitrate probably will crystallize out. (1)
5. The change in solubility between 0°C and 100°C is less for sodium chloride than for any of the three other compounds included in these data. (1)

This example requires understanding of the data before relating them to a line plotted on a graph.

Newton's law of gravitation is expressed by the equation $F = G \frac{m_1 m_2}{d^2}$

Where F represents the force between two masses, m_1 and m_2; d is the distance between their centers of mass; and G is a constant. If all other factors are held constant, which of the following graphs best express(es) the relation of F to d?



Analyzing an Unusual Problem

Stimulation of interest in the bizarre, the spectacular, or the fantastic to the neglect of the basic and fundamental aspects of the science course is certainly not to be recommended. However, it is sometimes possible to arouse interest by applying what has been learned to a problem situation that has a novel aspect. Some high school biology textbooks take the students far enough into heredity to involve them in problems concerning the inheritance patterns of certain human traits, such as the XX-XY pattern of sex inheritance, the inheritance of such sex-linked traits as color blindness and hemophilia, and the inheritance of the sex-influenced trait of baldness. The following set of four items on the inheritance of baldness is based on a situation which may serve to add some interest and should challenge the ingenuity of the better students.

EXAMPLE

Items 1-4 are based on the following situation which involves the inheritance of baldness, a sex-influenced trait.

A young female circus acrobat who hangs by her hair as a part of her act is wondering whether she should change her profession, if necessary, before it becomes too late. Her problem is this: Her mother is bald, but her father has a normal head of hair. Her older brother is rapidly losing his hair and will soon be bald.

Let B represent the gene for baldness and b the gene for non-baldness. In the heterozygous condition B is dominant in males, but b is dominant in females. A heterozygous man will be bald, but a heterozygous woman will not be bald.

1. Which of the following matings represents the parents of the young female circus acrobat?
   1. BbXX x BbXY
   2. bbXX x BBXY
   3. BbXX x bbXY
   *4. BBXX x bbXY
   5. BBXX x BbXY
2. The genotype of the older brother of the young female circus acrobat is represented by
   *1. BbXY
   2. bbXY
   3. BbXX
   4. BBXY
   5. BBXX
3. The genotype of the young female circus acrobat is represented by
   1. BbXY
   2. BBXY
   3. bbXX
   4. BBXX
   *5. BbXX
4. On the basis of the data, which of the following suggestions to the young female circus acrobat would be most justifiable?
   *1. You need not change your profession—according to the genetic evidence there is no likelihood at all that you will become bald.
   2. The chances are 1 in 4 that you will become bald.
   3. The chances are 3 in 4 that you will become bald.
   4. There is a 50% chance that you will become bald.
   5. Change your profession—according to the genetic evidence you are almost certain to become bald.

An Open-Ended Problem

We are sometimes faced with a problem for which the answer cannot be worked out from the available information. A proposed solution can, however, evoke discussion that will reveal whether students understand certain aspects of science upon which such a solution hinges. The following account which appeared in *Time* embodies such a problem situation.

The Bell of Kamela

Year after year, men cruising timber or hunting deer in the Blue Mountains of eastern Oregon had come back with the same story. Near the little hamlet of Kamela, they had often heard a faraway tinkling, a ghostly bell ringing. No one was ever able to track down the strange sound. It would fade away in the sighs of the wind through the big pines. Skeptics accused the men of hearing things.

Last week, slashing a right-of-way for a power line from Bonneville Dam, lumberjacks brought down a ponderosa pine. Tied by a shriveled leather thong, high in the tree top was the answer to the mystery of Kamela: a bronze cattle bell, inscribed with the date 1878. It carried the words, **Saignelégier—Chiantel—Fondeur.** Its clapper was worn smooth by years of gentle tinkling. The people of Kamela guessed that a pioneer had tied it to a sapling that grew into a towering pine."

PROBLEM: How did the bell get up into the tree?
HYPOTHESIS: (Suggested by the people of Kamela) A pioneer had tied it to a sapling that grew into a towering pine.

Students could be given this situation when they have completed their unit on plant structure and plant growth in a biology or botany class. They should be allowed some time to think about the situation. The statement of the problem and of the hypothesis could be omitted and a completely open-ended question—"What is wrong with this story?"—could be asked. Or "What is the problem in this situation? What proposed solution to the problem is offered in the selection? What is your appraisal of the suggested solution—do you agree or disagree with it? Why?" If the students are having difficulty thinking of sound reasons to support the stand they take, further questions such as the following could be put to them: "Have you ever seen a clothesline or a hammock fastened to a hook screwed into a tree trunk? Is it necessary to lower the hook a little each year? Have you ever seen a fence nailed to a tree? Does the tree carry the fence upwards with itself as the tree grows? Does a tree grow from the bottom up? If not, how does a tree grow? Is it likely that anything tied to the top of a very young tree will always in years to come still be in the top of the tree as the tree grows? Why or why not? At this point how do you appraise the hypothesis offered by the people of Kamela in solution to the problem as to how the bell got up into the tree?"

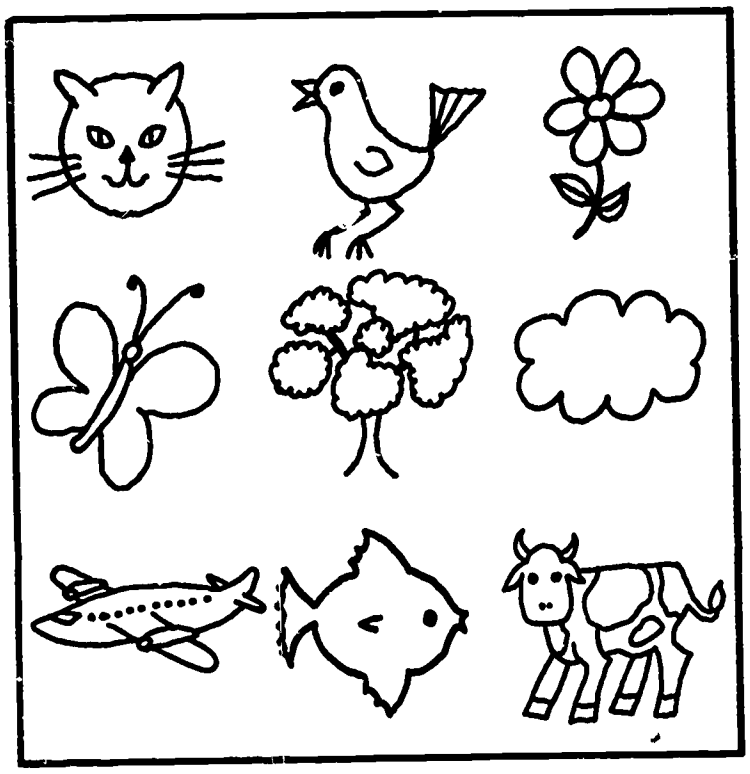
Even after the evidence in refutation of the hypothesis has been thoroughly considered and the hypothesis has been rejected, the problem still remains unsolved, and there is probably insufficient information given in the situation to serve as a basis for a solution. For this reason the situation lends itself best either to class discussion or for use in an unstructured test question which permits the student to write his own analysis of the story without necessarily presenting a solution to the problem.

III. APPLYING KNOWLEDGE TO CONCEPTS AT VARIOUS GRADE LEVELS

Test Exercises for Kindergarten

Science instruction at this level may take the form of teaching the children contrasting aspects of certain concepts. They should soon be able to recognize which alternative of a pair of opposites is being demonstrated. Perhaps one set of concepts is enough to introduce in one day's activity. This pair should be demonstrated and several examples shown, asking the pupils to identify which alternative is being considered in each case. The goal is attainment of an *understanding* of the concept rather than memorization of procedures. From the following test exercises, the nature of the presentation will become apparent. The answer sheet can be designed by the teacher in various ways, such as simple drawings that are related to the concepts. Since the drawings are to serve primarily as answer sheets, the outline figures should be made fairly small.

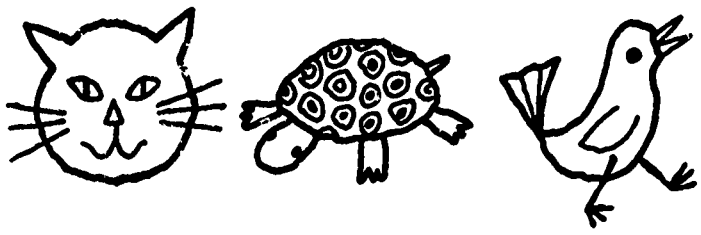
Picture tests, where a child puts an X on the picture of his choice, are simple to prepare. A picture card, such as the one in Item I, could be used in several ways to test concepts at different times. For example:



1. Put an X on the pictures of things that are **alive.**
2. Put an X on the picture of each thing that is an **animal.**
3. Put an X on the picture of each thing that is a **mammal.**

EXAMPLE

1. One of these three pictures is a kitty. Put an X on that picture.



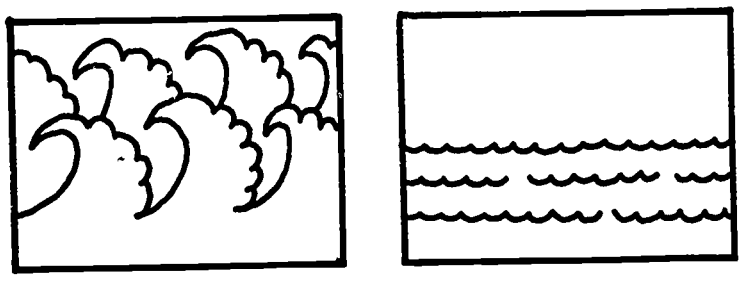
2. One of these three pictures shows what happens when something is heated or becomes warm. Put an X on that picture.

The teacher may wish to teach contrasting concepts, such as *rough* and *smooth*. After having taught the children the difference by using several examples, the same may be appropriate to test the children's understanding on a subsequent day, first by reviewing, then by presenting the test situation as follows:

The teacher holds up a piece of smooth board or a smooth rule, or any other smooth object, preferably of wood, but some other very smooth flat object may be used if more convenient. Running her hand over it, and also asking the children to do the same, she asks, "Is this rough or smooth?"

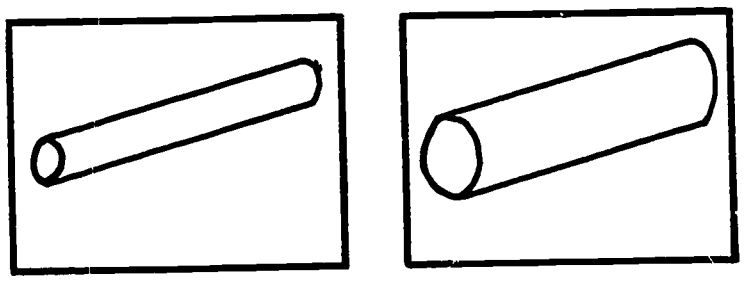
Then the teacher holds up a small piece of cordwood that still has the bark on it, or some other object that is very rough. She slides her hand over this object and allows the pupils to do the same, then asks, "Is this rough or smooth?"

Then for the test situation the teacher holds up a third object, perhaps a piece of very coarse sandpaper—the coarsest available—and asks the children to slide their hands over it. She says, "Put an X on the picture that is like this."



For another day's activity perhaps the contrasting concepts *thick* and *thin* might be presented and demonstrated. On a subsequent day the following test situation of the thick-thin concept may be undertaken:

The teacher holds up two books, one of which is very thick, and the other of which is very thin. She says, "One of these books is thick and the other one is thin." Then she asks the children to identify which book is which. She then puts these books out of sight and picks up a third book, which resembles one of the others in being very thick or very thin. She says, "Put an X on the picture that is like this."

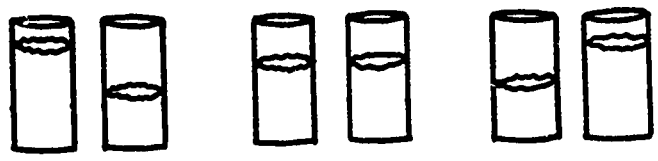


Test Exercises for First Grade

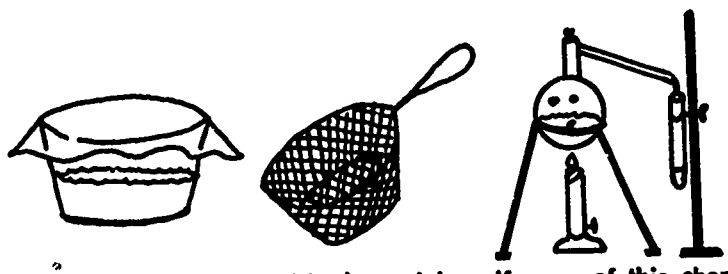
This set of questions will involve the use of a three-response answer sheet. The symbols should relate directly to the concepts the pupils are learning to understand. The teacher can read the question slowly and the students will mark an X on their answer sheets.

1. After some discussion of the characteristics of water and ice, the teacher may let several of the pupils fill small milk bottles, or other bottles that can be capped tightly, brimming full of water. The lids are placed on the bottles. If there is a refrigerator with a freezing compartment in the building, the bottles of water may be placed there to freeze until the next day. If it is freezing weather, the bottles could be placed outdoors in a safe place. Put each bottle in a transparent plastic bag fastened with a rubber band, for safety in case the bottle breaks. Usually the bottles will crack and some of the ice may expand considerably over the rim.

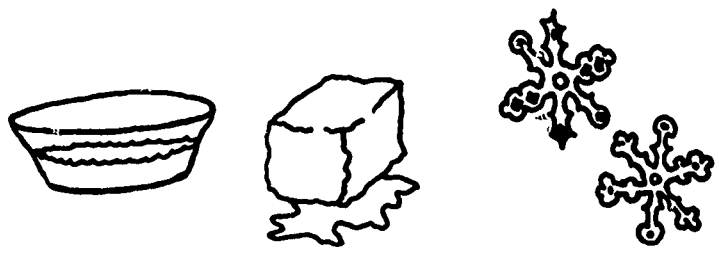
As a test exercise on the above, the teacher may hold up one of the cracked bottles with ice extending over the rim and ask why the bottle cracked when the water froze. When water freezes it takes up:



2. The teacher demonstrates boiling, filtering, and straining of water. The teacher then reads the test question. Pure water can be taken out of salt water by:



3. Salt has been mixed with chopped ice. If some of this chopped ice is packed around a container full of water, what will probably happen to the water in the container?



Test Exercises for Second Grade

Second graders may be able to record their responses by placing an X in a square. If this is so, the teacher could make an answer sheet with ten rows, each containing 3 squares, perhaps lettered a, b, and c, and make copies on the duplicator. She may read the questions to the class slowly, followed by the responses. The responses may have to be read several times.

1. A certain kind of clothes hook has a rubber cup on it. When the cup is pressed firmly against the wall or other smooth surface, the hook stays there. What keeps the hook in place? (The teacher may demonstrate this.)
   a. The stickiness of the rubber.
   b. Suction inside the rubber cup.
   * c. The pressure of the air in the room.

2. Jack weighs 56 pounds and Gail weighs 44 pounds. When they get on a seesaw, how should they sit to make it balance?
   * a. Jack should sit closer to the center.
   b. Gail should sit closer to the center.
   c. Each should sit the same distance from the center.

Test Exercises for Third Grade

Third graders may also use the same kind of answer sheet as was suggested for the second grade, or they may circle letters a, b, and c. The questions may be handed out in duplicated form, but the teacher should read the questions with the children and explain how to mark the answer sheets, unless they will be using the tests themselves as answer sheets, in which case they could underline or circle the correct answers.

1. When a watch is laid flat on a table, if 12 on the watch dial were to represent north, then 9 on the dial would represent
   a. south
   b. east
   * c. west
2. If 12 on your watch dial represents north, then what would represent southeast?
   * a. 4:30
   b. 7:30
   c. 10:30
3. A person standing at the seashore looking at a ship several miles from the shore can see the
   a. entire ship
   * b. upper part of the ship only
   c. lower part of the ship only
4. The horizon would be farthest away
   a. if you were standing on the seashore and looking out over the ocean.
   b. if you were on top of the Empire State Building or the Washington Monument and looking straight ahead.
   * c. if you were looking out of an airplane window while flying four miles above Earth.

Test Exercises for the Upper Elementary Grades

These children may be able to use a simple four-choice (need not be four) answer sheet that can be made on the typewriter or drawn by hand on a ditto master, thus

	a	b	c	d
1.	☐	☐	☐	☐
2.	☐	☐	☐	☐

The following exercises involve the analysis of a real-life situation which has scientific implications.

Steve, Dave, and Susie went on a picnic with their families. Soon they had a fire started and began to roast hot dogs. Someone had brought a jar of relish but could not unscrew the metal lid.

After several people tried to get the lid off, with no success, Steve said, "There must be some scientific principle that would apply here. Didn't we learn in science class that metals expand when they are heated? Glass expands but it does not get hot as fast as metal. So if metal and glass are heated together, the metal would expand more. If I could heat the metal lid carefully over a hot ember without warming the glass too much, the lid should expand before the glass does, and should turn easily."

Steve held the jar upside down over a hot ember, being careful not to hold the metal lid so close that he would be burned.

After warming the lid for about a minute, he wrapped a towel around it. When he tried to unscrew the lid it came off without any trouble.

For numbers 1-5 decide which of these items applies in each case.

a. Something Steve already knew (Fact)
b. Something that puzzled Steve (Problem)
c. Something Steve could figure out from what he knew (Deduction)
d. The theory on which Steve's success was based (Theory)

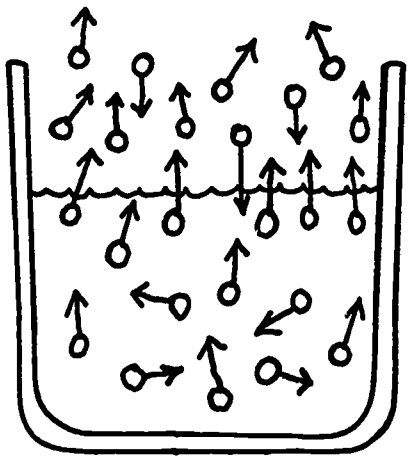
1. How could a simple scientific principle be used to get the metal lid to unscrew easily from the glass jar? (b)
2. Metals expand when they are heated. (a)
3. Glass also expands when it is heated. (a)
4. All things are made up of molecules. These molecules move faster when heated, causing the things being heated to expand. (b)
5. If Steve heated the lid without heating the jar, the lid would expand more than the glass and could be unscrewed. (c)

Test Exercises for Junior High School Level

Junior high school students can use commercially available regular answer sheets. If the local school does not have these, a simple answer sheet resembling the one suggested for upper elementary grades can be made.

EXAMPLE

For the following questions, refer to the diagram, which represents a container with liquid in it. The symbols are for molecules of the liquid and their direction of movement.



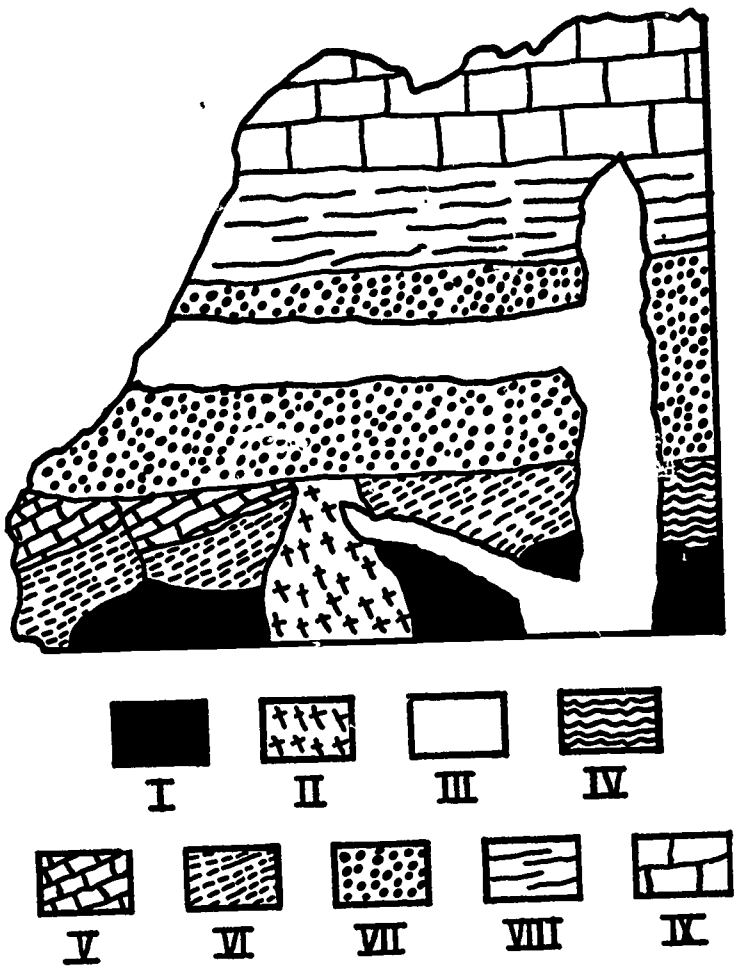
1. What is happening to the liquid?
   a. It is freezing.
   b. It is evaporating.
   c. The temperature is decreasing.
   * d. Both b and c above
   e. Nothing is happening to it.
2. One way of slowing the escape of the molecules would be to
   a. heat the liquid.
   * b. cool the liquid.
   c. blow air over the liquid.
   d. decrease the pressure of the air above the liquid.
   e. add more hot liquid to the container.
3. If a lid were to be placed on the container, which one of the following would be likely to occur?
   a. All of the liquid molecules would move into the liquid with none in the space between the liquid surface and the lid.
   b. All of the air molecules would move into the spaces left by liquid molecules in the liquid itself.
   c. The temperature of the liquid would decrease.
   * d. The number of liquid molecules leaving the liquid would be approximately equal to the number of liquid molecules entering the liquid.
   e. All of the above would be likely to occur.

4. Which one of the following would be a way to increase the rate at which the molecules could escape from the liquid?
   * a. Blow air over the top of the container.
   b. Put a lid over the container.
   c. Cool the liquid.
   d. Pour oil over the liquid.
   e. Increase the air pressure over the surface of the liquid.

Interpretation of a land-feature diagram that has not been studied in class can serve to elicit responses from the students which reveal how well they understand the concepts they have learned.

EXAMPLE

Items 1-5 refer to the cross-sectional land feature diagram. Formations are indicated by Roman numerals.



1. Fossils would be least likely to occur in
   * a. III b. V c. VII
   d. VIII e. IX
2. An unconformity exists between
   a. I and III
   b. III and VI
   * c. V and VII
   d. VII and VIII
   e. VIII and IX
3. The youngest formation is
   a. I b. II c. III d. V * e. IX
4. The oldest formation is
   * a. I b. II c. III d. V e. IX
5. Which formation is made up of igneous rock?
   * a. III
   b. V
   c. VII
   d. VIII
   e. IX

Similarly, the student's interpretation of a *chemical formula* that he has not previously encountered can reveal whether or not he understands the symbols and conventions he has studied in class.

1. The number of atoms of oxygen in the formula $2Al_2(SO_4)_3$ is
   a. 4
   b. 7
   c. 8
   d. 12
   * e. 24

Test Exercises for Senior High Level

On occasion it may be desirable to have the students solve a problem delineated in the first of two paired items and to give the basis for the solution in the second item of the pair. The following paired items from high school chemistry tests illustrate this format. The odd-numbered items present problems to be solved. The even-numbered items that follow require determination of which law(s) or principle(s) is/are needed to solve the previous odd-numbered problem.

Laws and Principles

I. Equal volumes of gases at the same temperature and pressure contain equal numbers of molecules.

II. At constant temperature, the volume of a given mass of gas is inversely proportional to the pressure exerted by the gas.

III. At constant pressure, the volume of a given mass of gas is directly proportional to its absolute temperature.

IV. Every pure chemical compound always contains definite, constant proportions or relative masses of the elements of which it is composed.

V. A law or principle not listed above

1. A certain gas occupies a volume of 10 liters at STP. What volume would this gas occupy at 380 mm pressure?
   a. ½ liter
   b. 5 liters
   c. 10 liters
   * d. 20 liters
   e. 380 liters
2. What is the basis for the solution? (Select from laws and principles listed above.)
   a. I only
   * b. II only
   c. III only
   d. II and III
   e. II and V
3. If the atomic mass of carbon is 12, and that of oxygen is 16, the mass of carbon dioxide which could be produced from the complete oxidation of 12 grams of carbon ($C + O_2 \rightarrow CO_2$) is
   a. 12 grams
   b. 16 grams
   c. 28 grams
   * d. 44 grams
   e. none of the above
4. What is the basis for the solution?
   a. I only
   b. I and III
   c. I and IV
   * d. IV and V
   e. V only

IV. SELECTING AND PREPARING READING PASSAGES FOR TESTING SITUATIONS

Interest in present-day scientific activity, while it should not completely obscure and crowd out all consideration of classical experiments of the past, ought to be encouraged by every means possible. Every week the leading news magazines report on current scientific activity. Some of these reports can be read and understood by high school students. Perhaps even better sources of such reporting are the specially prepared digests of scientific activity such as *Science News.* These digests often present very abbreviated accounts in which most of the data are omitted. For items of special interest, however, the original source, e.g., *Science* or *The Journal of the American Medical Association,* can be located and the complete data obtained. While the details may be complex, the methodology employed in these experiments is sometimes not too difficult for students to understand. Articles of unusual and timely interest can be partially rewritten or paraphrased by the teacher into somewhat simpler language, while retaining the fundamental ideas. If each sentence in the article is copied as a separate paragraph, it is easy to see which ones contribute most to the discussion and which ones can be shortened or combined.

For a test, students may be asked to make an analysis of one or more paragraphs to locate the central problem, suggested hypotheses, assumptions implicit in the situation, false assumptions, evidence which supports a particular hypothesis, evidence which refutes a given hypothesis, illogical conclusions, irrelevant arguments, and the like.

Articles dealing with scientific or science-related activity which might serve as the nucleus for development of thought questions can sometimes be found in newspapers, such as the Sunday *New York Times,* or in news magazines such as *Time, Newsweek,* and *Reader's Digest,* or in digests of scientific activity such as *Science News* or *Science Digest.* These articles, however, have usually been rewritten from original research papers that were presented at recent scientific meetings or which have been published in current issues of scientific journals.

The most useful purpose served by the newspaper or news magazine account is to announce a new discovery, process, product, or theory in a place where the general reader is apt to come across it. However, many news reports of scientific activity are not suitable for evaluation in the form in which they appear. This may be due to: (1) The conclusion is often stated in exaggerated or misleading fashion in the title or in the first sentence of the article, or (2) the article may be purely descriptive with no reasoning presented, or (3) data which can be analyzed for evidence may be entirely lacking, with the result that there is no basis for judging for oneself the validity of the conclusion or the manner in which the results were obtained. The teacher of science can then trace the information to the original source and there obtain the data which are required for the construction of questions requiring thought and reasoning.

Examining Readings for Validity

Gerald Wendt, former science editor for *Time* and *Science Illustrated,* suggests the following criteria as guides to judging the validity of information presented in a newspaper or magazine article. (12)

For news publications, according to Dr. Wendt, reliability is favored if:

1. The publication has a general reputation for careful, complete, and accurate reporting of all news.
2. The story is derived from an announcement at a professional scientific meeting or in a professional scientific journal.
3. The investigation was conducted at a recognized university or research center, rather than by an individual working on his own.
4. The report is based on evidence derived from repeated experiments or confirmed by separate methods, rather than being based on observation of a single case or event.
5. The headline is reasonably consistent with the story.
6. The first paragraph of the story is free of lurid statements and unsupported assertions, and is consistent with the rest of the story.
7. Effort is shown to report the *real* story by the use of quotations or by reference to the original investigator and his work.
8. The story is long enough to show that it was carefully prepared.
9. The full story is reproduced and not cut to fit available space. The last paragraph is likely to be valuable as a summary.
10. The story is free from: (a) excessive extrapolation; (b) phrases like "which may," used merely to create interest.

For entertainment magazines, again, according to Dr. Wendt, reliability is favored if:

1. The periodical is well established and has a reputation for general reliability in its nonfiction articles.
2. The author of the article
   a. is one of the scientists who participated in the investigation mentioned in the article, or
   b. is a scientist who is active in the field with which the article deals, or
   c. indicates clearly how he obtained the information and ideas in the article.
3. The diagrams, pictures, and their captions are consistent with the ideas developed in the article.

4. The author, in presenting only one point of view, acknowledges this and makes clear what the other viewpoints are.
5. The author makes clear distinctions between factual information and opinions or hypotheses.
6. The author shows an attitude of tentativeness and objectivity and takes care to support the inferences that he makes.
7. The author avoids reasoning by analogy or extrapolation.
8. The author devotes a fair share of the article to explaining *how* the conclusions were obtained rather than devoting most of the article to elaborating and interpreting the conclusions.

Probably very few news stories or magazine articles will be found that satisfy all the conditions included in the above list, and they may fulfill a condition only partially. However, the conditions listed will probably suffice as a basis for evaluation that avoids the extremes of gullibility on the one hand and cynicism on the other.

An atmosphere of gentle skepticism should prevail until authority is established or until independent proof has been reported.

A passage from a book may prove to be fruitful as a basis for constructing test items. The following outline, adapted from an unpublished document by Professor Merle Harris of the University of Minnesota, suggests methods that may be of value in developing test items that relate to such a published article or excerpt from a book:

I. If a problem (or theory) is involved in the article, questions might bring out whether the student
   A. understands the problem or theory.
   B. believes the problem is solved or the theory verified:
      1. If so, what evidence does he regard as clinching the argument?
      2. If not, what evidence does he regard as being needed to clinch the argument?
         a. Does he believe the needed evidence is obtainable now? Later? Never?
   C. sees the relation of the problem or theory to other phenomena.
   D. understands the assumptions made when the problem was posed or the theory stated.
   E. understands further assumptions made when the solution was given.

II. If a paradox is involved, questions may be asked to determine
   A. what the student's attitude toward it is. That is, should the intelligent layman
      1. attempt to resolve the paradox on the basis of his knowledge of the principles involved?
      2. not attempt to resolve it because
         a. it should be left to the experts?
         b. it is impossible even for the experts?
   B. whether the student understands
      1. what kind of evidence or investigation is needed to resolve the paradox.
      2. the possibility of obtaining the evidence needed to resolve the paradox.
      3. the assumption which, if true, would resolve the paradox (if evidence is unobtainable).
      4. the consequences of resolving the paradox.

III. If the article describes a process, questions may be asked to determine whether the student
   A. understands the process in the sense that he
      1. recognizes the scientific principles involved.
      2. sees the relationship between this process and similar processes.
   B. recognizes that the process depends on sound scientific principles.
   C. recognizes that the process is sufficiently developed to accomplish some stated purpose.

IV. If the article describes a relationship, questions may be asked to determine whether the student
   A. believes the evidence presented is
      1. sufficient to establish the relationship as correct.
      2. sufficient to indicate that the relationship is probably correct.
      3. insufficient to give any indication either way.
      4. sufficient to cast doubt on the truth of the relationship.
      5. sufficient to prove the relationship false.
   B. sees the significance of the relationship, if true.

V. If the article describes a scientific experiment or discovery, questions may be asked to determine whether the student
   A. understands precisely *what* is new in the experiment or discovery.
   B. regards the experiment or discovery as having been investigated
      1. properly thus far, but warrants extension.
      2. properly thus far, and might well end at this point.
      3. improperly because certain necessary steps have been carelessly done.
   C. believes the results have been properly interpreted.
   D. believes the results have been improperly interpreted.
   E. sees the impact of the interpreted results on previous assumptions in this field.

BIBLIOGRAPHY

1. Adkins, Dorothy. *Test Construction.* Charles E. Merrill Books, Inc., Columbus, Ohio. 1960.
2. Balch, John. "The Influence of the Evaluating Instrument on Students' Learning." *American Educational Research Journal* I: 169-182; May 1964.
3. Bloom, Benjamin, and Krathwohl, D. R. *Taxonomy of Educational Objectives, Handbook I:* "Cognitive Domain." David McKay Co., Inc., New York. 1956.
4. Bradfield, James M., and Moredock, H. Stewart. *Measurement and Evaluation in Education.* The Macmillan Company, New York. 1957. p. 204.
5. CUEBS Panel on Evaluation and Testing. *Evaluation in the Biological Sciences.* Commission on Undergraduate Education in the Biological Sciences, Washington, D.C. 1967. (In press)
6. Dressel, Paul L., *et al.* "How The Individual Learns Science." In *Rethinking Science Education.* Fifty-Ninth Yearbook, Part I. National Society for the Study of Education. University of Chicago Press, Chicago, Illinois. 1960. pp. 39-62.
7. Ebel, Robert L. *Measuring Educational Achievement.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey. 1965.
8. Hedges, William D. *Testing and Evaluation in the Sciences for the Secondary Schools.* Wadsworth Publishing Company, Belmont, California. 1966.
9. Heil, Louis M., *et al.* "The Measurement of Understanding in Science." In *The Measurement of Understanding.* Forty-Fifth Yearbook, Part I. National Society for the Study of Education. University of Chicago Press, Chicago, Illinois. 1946. pp. 104-137.
10. Holton, Gerald. "Testing and Self-Discovery." *University College Quarterly* 9: 3-9; November 1963. Michigan State University, East Lansing, Michigan.
11. Kessen, William. "Statement of Purposes and Objectives of Science Education in School." *Journal of Research in Science Teaching* II (1): 3-6; 1964.
12. Klinckmann, Evelyn. "The BSCS Grid for Test Analysis." *BSCS Newsletter* 19: 17-21; September 1963. Biological Sciences Curriculum Study, University of Colorado, Boulder.
13. Lucow, William Harrison. "The Whole Truth and Nothing but the Truth Examination." *Science Education* 41: 99-103; March 1957.

    Description of a test designed to defeat the efforts of the "test-wise" opportunist type of student and to encourage the student who is conscientious and thorough. The rationale for the examination is presented together with illustrative items. This discussion is followed by a 30-item test and key to the answers.
14. Mead, Margaret, and Metraux, Rhoda. "Image of the Scientist among High School Students." *Science* 126: 384-390; August 30, 1957.

    Excellent recommendations for more vitalized science teaching to offset the negative attitudes toward science are set forth in this report.
15. Mason, John M. "The Direct Teaching of Critical Thinking in Grades Four Through Six." *Journal of Research in Science Teaching* I (4): 319-328; 1963.
16. Nelson, Clarence H., and Mason, John M. "A Test of Science Comprehension for Upper Elementary Grades." *Science Education* 47: 319-330; October 1963.
17. Nelson, Clarence H. "Evaluation of the Objectives of Science Teaching." *Science Education* 43: 20-27; February 1959.
18. Nelson, Clarence H. "Evaluation in the Natural Sciences." In *Evaluation in Higher Education* by Dressel, Paul L., *et al.* Houghton Mifflin Company, Boston, Massachusetts. 1961.
19. Noll, Victor H. *Introduction to Educational Measurement.* Houghton Mifflin Company, Boston, Massachusetts. 1965.
20. Noll, Victor H. *Science Education in American Schools.* Forty-Sixth Yearbook, Part I. National Society for the Study of Education, University of Chicago Press, Chicago, Illinois. 1947.
21. Novak, Alfred. "Scientific Inquiry." *BioScience* 14: 25-28; October 1964.
22. Palmer, O. E. "Rater Reliability on Themes." In *Evaluation in the Basic College* by Dressel, Paul L. *et al.* Harper and Brothers, New York. 1958. pp. 127-128.
23. Pettit, Lincoln C. *How to Study and Take Exams.* John F. Rider, Inc., 116 W. 14th St., New York. 1961.
24. "Special Feature on Testing." *NEA Journal* 52: 16-25; October 1963.
25. Tannenbaum, Harold E.; Stillman, Nathan; and Piltz, Albert. *Evaluation in Elementary School Science.* Circular 757. U.S. Office of Education, Washington, D.C. 1964.
26. Walbesser, Henry H. "Curriculum Evaluation by Means of Behavioral Objectives." *Journal of Research in Science Teaching* I (4): 296-301; 1963.

    Walbesser, in working with the AAAS elementary school project entitled SCIENCE—A Process Approach, evaluates pupil performance individually. Each pupil is asked to demonstrate with simple apparatus how he or she would go about solving a given problem situation.

REMINDERS

1. Determine the objectives or goals of the course.
2. Set up a list of content headings.
3. Construct a two-axis chart or grid with content headings listed on one axis and objectives listed on the other.
4. Make an apportionment of items to each content heading and to each objective.
5. Determine what shall be the basis for the item writing, e.g., textbook laboratory guide, reference books, field situations, reading selections from news magazines or journals, films, or other sources.
6. Have other science teachers take the completed test.
7. Save some of the best questions for later tests.

Add your own reminders below.