A Study of Type Questions for General Science Tests

by

Donald McIntosh Flather

A Thesis Submitted for the Degree of Master of Arts in the Department of Education

The University of British Columbia
April, 1939.
ACKNOWLEDGEMENTS

The writer is indebted very deeply to Professor C.B. Wood of the Department of Education, University of British Columbia, for his continual assistance, inspiration, and constructive criticism as well as for his permission to submit the questionnaire to his class in Education 23 during the summer of 1938.

The writer wishes to thank Dr. Wm. Black, also of the Department of Education, University of British Columbia, for his lucid general criticism of the thesis and for his many valuable suggestions that have improved this report.

Further assistance has been obtained from Mr. Lord, Principal of the Provincial Normal School at Vancouver who assisted in submitting the questionnaire to the members of his class last summer, Education 12.

Constructive criticisms have been received also from Dr. J.B.W. Pilcher (formerly of U.B.C.), Dr. H.B. King (Technical Adviser to the Department of Education, Victoria), Mr. Robt. Straight (Bureau of Measurements, Vancouver School Board), and from Dr. W.J. Osburn (Department of Education, University of Washington).

D.M.F.
This research was undertaken to demonstrate what could be done in the way of testing the achievements and the growth of abilities which form the anticipated outcomes of the courses in the high schools, General Science IV and V.

As the research has turned out it is really only an introduction to the problem, for many subsidiary problems that were uncovered have been left untouched. The writer had hoped to make a complete survey, but this desire had to be narrowed down considerably.

Further detailed study of the testing procedures as applied to each objective of the courses is needed, as is also a more scientific way of selecting the objectives of a course and then evaluating them. In fact, the limitation of research to any one of the several objectives is really a major problem itself.

While the work may not be as complete as it might have been, there is a considerable amount of information presented that can be used in a programme of testing in General Science. The writer has profited greatly by the research, and has been able to apply many of these suggestions found in the text to his own programme of testing.
# TABLE OF CONTENTS

## Chapter I  AN OUTLINE OF THE PROBLEM
General Conditions Affecting the Problem  1
Local Conditions Affecting the Problem  2
Dearth of Recent Experiments on Testing  5
Summary  7

## Chapter II  AN EXAMINATION OF THE PRESENT OBJECTIVES OF GENERAL SCIENCE IV AND V
Scope of Chapter  8
The Objectives of General Science IV and V  8
Examination and Analysis of Present Objectives  9
Evaluating the Importance of the Present Objectives  18
Re-arrangement of the Present Objectives  29

## Chapter III  COMMON FORMS OF TEST ITEMS AND THEIR FUNCTIONS
Scope of Chapter  31
The Three Major Types of Tests  31
Research Limited to Questions Testing Achievements  34
Chief Forms of Questions in Common Use  35
The True-False Question  36
The Multiple-Choice Question  38
The Completion or Recall Question  40
Matching Questions  41
The Analogy Question  42
Diagrams  44
The "Catechism" Type of Question  45
The Essay Type of Question  46
Mathematical Calculations  53
Checking Lists, and "Identifications"  54
The Performance or "Practical" Test  54
Oral Examinations, Interviews  56
Re-arrangements  58
Comparisons and Contrasts  58
Constructions and Drawings  59
Comments on Reliability  60
General Conclusions  61

## Chapter IV  A STUDY OF PRESENT STANDARDIZED TESTS IN SCIENCE
Purpose of This Study  62
Tables Summarizing Investigations  63
Highlights of Tests Examined  79
Applicability of Standardized Tests to Testing Programme  80
Specimen Test Questions from the Standardized Tests  83
INDEX TO TABLES

<table>
<thead>
<tr>
<th>Table I</th>
<th>Summary of Returns of Questionnaire Submitted to Two Classes of Teachers, Summer Session, U.B.C., 1938</th>
<th>page 22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table II</td>
<td>Evaluations and Weightings Made by Sixteen Administrators, Principals, and Science Committee Members</td>
<td>25</td>
</tr>
<tr>
<td>Table III</td>
<td>Summary of Evaluations by 78 Teachers, U.B.C.</td>
<td>26</td>
</tr>
<tr>
<td>Table IV</td>
<td>Analyses of Standardized Chemistry Tests</td>
<td>64</td>
</tr>
<tr>
<td>Table V</td>
<td>Analyses of Standardized Physics Tests</td>
<td>68</td>
</tr>
<tr>
<td>Table VI</td>
<td>Analyses of Standardized Biology Tests</td>
<td>70</td>
</tr>
<tr>
<td>Table VII</td>
<td>Analyses of Standardized General Science Tests</td>
<td>72</td>
</tr>
<tr>
<td>Table VIII</td>
<td>Analyses of Aptitude Tests and Others</td>
<td>74</td>
</tr>
<tr>
<td>Table IX</td>
<td>Analyses of Matriculation Papers</td>
<td>76</td>
</tr>
<tr>
<td>Table X</td>
<td>A Comparison of Frequency of Question Forms on Tests</td>
<td>79</td>
</tr>
<tr>
<td>Table XI</td>
<td>Abstract from Ruch and Rice, Frequency of Question Forms</td>
<td>79</td>
</tr>
<tr>
<td>Table XII</td>
<td>Comparison of Degree Which the 29 Standardized Tests Which Were Examined Measure the Achievements Other Than Knowledge</td>
<td>80</td>
</tr>
</tbody>
</table>

Questionnaire Blank

Appendix 1

INDEX TO FIGURES

| Figure I | Frequency Distribution of Rankings of the Ten Objectives by Fifty-two Secondary School Teachers | 23 |
| Figure II | Frequency Distribution of Seventy-eight Evaluations of Objectives | 27 |
| Figure III | Flower, (Ruch-Popence Fig. 2, Form A) | 83 |
| Figure IV | Lever, (Ruch-Popence Fig. 4, Form A) | 83 |
| Figure V | Pulleys (Ruch-Popence Fig. 10, Form A) | 85 |
| Figure VI | Pump (Ruch-Popence Fig. 15, Form A) | 84 |
| Figure VII | Photosynthesis (Ruch-Popence Fig. 16, Form B) | 84 |
| Figure VIII | Collecting Carbon Dioxide (Ruch-Popence, Fig. 10, Form B) | 85 |
| Figure IX | Bunsen Burner Flame (Persing Fig. 6, Form A) | 85 |
| Figure X | Collecting Gases over Water (Persing) | 85 |
| Figure XI | Gas Generating and Collecting (Persing Form A) | 87 |
| Figure XII | Graph, Analysis of Copper Sulphide | 87 |
| Figure XIII | Illusions and Accuracy Test (Stanford Scientific Aptitude) | 88 |
| Figure XIV | Reading Weights on Toledo Scales | 116 |
| Figure XV | Beam Balance | 117 |
| Figure XVI | Rock Strata | 118 |
| Figure XVII | Graph, Air Resistance and Velocity of Car | 122 |
| Figure XVIII | Liquids in Graduated Measures | 173 |
| Figure XIX | Disassembled Apparatus | 173 |
| Figure XX | Clam | 174 |
| Figure XXI | Flower | 174 |
| Figure XXII | Stem of Tilia, Transverse Section | 175 |
| Figure XXIII | Relative Fields of Testing for Mastery, Administrative, and Achievement Tests | 196 |
The purpose of this research is to survey what has been done in developing the techniques of testing in the field of General Science, and in other fields that might make contributions, and to show how the forms of tests and of questions may be applied to measure progress of pupils towards the objectives of General Science courses.

General Conditions Affecting the Problem.

With the broadening of the conception of education that has occurred in the last twenty-five years there has been an increase in number of the objectives of any course in the secondary school. At the turn of the century the disciplinary value of a course was the chief objective, and this in turn was followed by what might be termed the "subject matter era" wherein the greatest emphasis was laid on the accumulation of facts. Science courses have gone through this transition like all others. When the restraining bonds were broken at last a period developed in which educators seemed to outdo themselves in setting up objectives, even by the hundreds for some courses. All these objectives obviously could not be attained because it was humanly impossible. This period of excessive expansion of objectives has given place to a more moderate selection of attainable objectives at the present time.

With the growing emphasis on a moderate range of attainable objectives for any good science course a swing away from the narrow informational or factual type has occurred. Concomitantly with this change in educational philosophy there should have been a marked parallelism of methods of testing achievements of the several objectives. Up to the
present time this movement has not been vigorous, for only occasional
evidence is presented of attempts to meet this new need of adapting meas-
uring devices to the objectives.

To study the functions and applicability of question forms is a use-
ful service at this stage of the development of science courses in general
because many are being revised. Further, to know how and when to use the
forms of questions is also of considerable importance to the average
teacher in General Science in order that he may measure definitely the par-
ticular achievement under examination.

This research may be of assistance also in helping a teacher to real-
ize that the testing of achievement of any objective must be done directly
and not by round-about methods. It should assist the teacher to be care-
ful in the choice or selection of ideas that he wishes to combine in the
preparation of questions and tests of a specific nature.

Local Conditions Affecting the Problem.

Besides these general reasons for the research there are several
local conditions that demand a survey of testing procedures in the field
of General Science. At the present time there is considerable bewilder-
ment among teachers of General Science in British Columbia concerning
testing procedures. This confusion is observable in discussions of the
Science Section of the Secondary Teachers' Association of the Lower Main-
land, in letters to the magazine "The B. C. Teacher", and in numerous
letters written to the Department of Education on the matter of testing.

There are four local or provincial factors that are responsible for
the confusion in testing in General Science. First of all, a new type
of science course has been projected into the upper grades of the high school in the form of General Science IV and V which first became operative in 1937 and 1938. These courses are decided innovations and because of this fact very little experience up to this time has been gained in the preparation of suitable achievement tests. The Department of Education has not yet given any guide to the teachers of science by the setting of a paper in General Science IV and V but it is taking steps to develop satisfactory examinations. The older tests on the special sciences can no longer be taken as criteria, and only the occasional test item on these papers is ever likely to be needed in testing in General Science. Further, testing procedures in the past in General Science have been in advance of those in the special sciences, and these procedures are having more effect in modifying the testing in the special sciences than the reverse case.

The second important factor is the attitude of the Department of Education which insists in accordance with the educational philosophy published in its programme of studies that the testing of pupil achievement in General Science must be done in agreement with the objectives of the course. The Department of Education has announced also that students who will write the first graduation paper in General Science in June 1939 will be held responsible for work done in General Science III, IV and V, and that teachers must align their own testing programmes with the objectives of the courses in question. The Department of Education has stated that "The examination papers for matriculation shall be in conformity with the principles laid down" by itself. In the eyes of the advisers to the

1. Programme of Studies for the Senior High Schools of B.C., Bulletin VIII, p. 60
2. Programme of Studies for the Senior High Schools of B.C., Bulletin 1, pp. 20-21
Department testing assumes a great importance and "must be regarded as an essential part of the teaching process". Continuing in this theme the Department adds that "It is fundamental that testing should bear upon the objectives of a course". This new emphasis upon testing is indicated still more clearly by the quotation that "Knowledge and skills are beyond question important, but are not the only outcomes of the educative process. An examination, or an examination system, which stresses these outcomes to the exclusion of the others, not only excludes these outcomes from the testing procedures, but speedily excludes them from the teaching procedures as well. It would seem quite clear from the foregoing statements that a fuller, more scientific programme of testing to include the hitherto untested phases (in British Columbia at least) should follow and keep pace with the ideals of the last revision of the science courses.

The third factor of importance comes to light in studying the problem of the accrediting of high schools. Such accrediting does not relieve the school in the slightest degree of its responsibility for establishing sound testing procedures, but rather increases the need of the accurate measuring of achievement. The school may be freed from the rigours of long hard examination periods but is not relieved from its duty to measure achievements in a precise manner. Further, because accrediting involves the University of British Columbia it is desirable that exact measurement of achievements be developed, for the positions of both the high school and the University must be protected. Accurate measurements of achievements would tend to prevent difficult situations from arising.

1. Programme of Studies, Bulletin 1, pp. 20-22
2. Programme of Studies, Bulletin 1, pp. 20-22
Finally, the whole educational structure in British Columbia is being criticized rigorously in certain quarters and attacked openly by many persons who attempt to keep down the costs of education. Many expenditures are attacked as being unnecessary, or as providing only the "frills" of education. Very frequently adequate science equipment is classified under this latter heading, as is also the building of suitable rooms in which to conduct science teaching. Exact evidence, if it can be obtained, of pupil achievements of the objectives of the courses in science is necessary to defend the cost, time, and effort spent in science training. If a course claims to develop more than knowledge achievements it must have some measure of these broader accomplishments. In order to support expenditures for equipment and buildings objective evidence of the development of scientific habits of thinking, of attitudes, of techniques, and of the other outcomes is highly desirable.

Dearth of Recent Experiments on Testing.

A consideration of the first four local factors mentioned would suggest that the time is now ripe for examining seriously what types of tests are applicable to General Science in order that pupil achievement of the various objectives of the science courses may be measured. In carrying out this idea the first step was to examine the literature on testing, and more particularly on science testing. There appeared to be relatively little on achievement testing except in the field of measuring factual matter. (With the latter field almost all teachers are quite familiar.) However, there are several unrelated reports of research yielding ideas that can be incorporated into a more extensive programme of testing designed to evaluate some of the more subtle aspects. The
incorporation of these contributions into the science and practice of educational measurement would be in itself a valuable service.

According to Dr. J. B. W. Pilcher there has been no significant contribution to the science of educational measurements in the last ten or twelve years, although refinements and repetition of original experiments have occurred. There is need at the present time of a re-examination of past contributions to measurements in the light of further experience and newer concepts of the function of teaching. This need was foreseen by Dr. A. W. Hurd in 1929 who concluded after an extensive survey of testing techniques in physics that "Existing measures of educational products need further study and improvement. Many abilities, presumably developed by science instruction, have not been measured objectively, and there is no objective proof that the claims (of developing these other abilities) are just." "Abilities presumably gained in the laboratory have not been measured objectively as yet, though a beginning has been made." A survey on the part of the writer revealed that comparatively little had been done along that line since Dr. Hurd made the foregoing statement. Not very much has developed that could be turned into immediate use in a testing programme devised in accordance with the objectives of General Science IV and V. This inadequacy of testing techniques in measuring pupils' growth along the lines of the objectives of the course provided a challenge to modify old or to invent new techniques.

Teachers claim, possibly with a great deal of truth, that the teaching of science has improved during the last fifteen years, particularly in the development of abilities other than the gaining of factual

2. Curtis, F.D. Second Digest of Investigations in the Teaching of Science
material. Present day tests rarely show this trend, not because it is not there, but because they are inadequate measuring instruments for the purpose. Although it may be desirable to develop adequate instruments to show this change it is more important that these tests be developed for regular use in the ordinary routine of teaching, so that all attainments made by pupils in the field of General Science may be judged.

Summary

In general the plan of the research was this; first, to read generally on the problem of testing particularly as it applies to science; secondly to examine the present objectives of General Science IV and V and to find some way of arranging the objectives in order of importance. The next major step was to analyze common types of questions to see if they are adaptable to any or all testing purposes. This step completed, the next logical move was to examine in detail present day standardized tests to see if one or a combination of several would serve the needs of the courses. If such were found naturally no further work would be needed. If none could be found, naturally the duty would devolve upon the writer of trying to improve or modify present forms or to invent new techniques. The results of this research will be summarized in the final chapter.
CHAPTER 11

AN EXAMINATION OF THE PRESENT OBJECTIVES OF

GENERAL SCIENCE IV and V

The Objectives of General Science IV and V.

At the present time the courses in General Science IV and V have ten objectives which were chosen by the revision committee after long and careful deliberation. These objectives were thought to be sound, reasonable, and attainable. Stated as they appear in the bulletin they are:

1. To acquire a body of knowledge in the field of science which will enable the student to interpret and appreciate his environment.

2. To develop ability in the use of the scientific method, e.g.:
   (a) To make accurate observations, and to record them systematically.
   (b) To draw valid conclusions.
   (c) To suspend judgment until sufficient evidence has been obtained.
   (d) To develop a critical yet tolerant attitude toward new ideas.

3. To develop the ability to perform simple experiments, and thus to appreciate the experimental basis of science.

4. To enable the student to counteract superstition and to correct erroneous beliefs through the application of scientific principles.

5. To appreciate achievements in the field of science, and the contributions of scientists to the modern world.

6. To explore the field of science in order to assist the pupil to choose his vocation.

7. To provide materials for a worthy use of leisure.

8. To develop the desire to read scientific literature.

1. Programme of Studies for the Senior High Schools, Bulletin 1, 1937, p.159
9. To develop resourcefulness and adaptability to new conditions.

10. To acquire knowledge which will contribute to public and personal health.

These are the objectives which will guide science teachers for the next few years. Whether a teacher thinks that they are too extensive or too restricted they are open to modification only by the proper authority, the Department of Education.

**Examination and Analysis of Present Objectives.**

In order that both teaching and testing in General Science may be most effective it is necessary to examine the ten objectives stated above in a very careful manner. Such an examination should include both a critical study of the implications of each objective together with a weighting or evaluating of each objective relative to the others of the set. Upon these two bases each objective will be analyzed in the chapter.

The need of an analysis of the objectives is evident in the frequency with which teachers question the implications of one or more of them. Few teachers agree exactly in their interpretations of objectives unless these are elaborated or qualified in some way. Because this task has not been done in the courses in question attention must be given it in this chapter.

The evaluation of objectives is advisable in order to obtain a standard, albeit a conservative one, to be used in stabilizing the teaching of a course. Almost every teacher has a bias toward some objective of a course, and in certain cases this is so strong as to subjugate the others or to exclude them completely. All teachers, either overtly or covertly, sift the objectives to find which they think the most important. A more
general measure on the value of objectives than the opinion of one person surely would be more valuable and reliable. Likewise the most important objectives will be demonstrated. These facts are needed in order to allocate the teaching effort, testing time, and credits. There seems no valid reason why the evaluation of the objectives should not be expressed in a composite opinion of those most qualified to give opinions, namely the teachers in the secondary schools, professors, and those in central positions in the educational system. To obtain this general measure a questionnaire was submitted to these persons.

Because the science revision committee appointed by the Department of Education did not make any attempt to evaluate the objectives, it became necessary to do that in this research. It is desirable to know not only what to teach but also how much of it to teach. The committee did not go this far in its deliberations but simply picked out what it considered to be the most suitable objectives from a rather extensive list selected from many sources. Thus the weeding-out process is a sort of crude evaluation of the ten objectives as against those not selected, but it is not accurate enough for a basis of testing. To assume equal value for each objective is unwarranted, and to assume that the objectives as listed in the Bulletin I are ranked in order would be wrong for the committee made no attempt to organize the objectives in order of value. The present order in the Bulletin therefore has no intentional significance.

The part immediately following will be devoted to an analysis of the objectives. This in turn will be succeeded by the evaluating of the objectives by means of the questionnaire. Each of the ten objectives will be dealt with in turn.
1. "To acquire a body of knowledge in the field of science which will enable the student to interpret and appreciate his environment". This objective is meant to include the factual side of science teaching, the scientific facts, the laws or principles, and material of a similar nature. In fact its field corresponds to the usual connotation given to "subject matter". In General Science IV and V there is to be a wide range of material expressing the basic principles in broad and simple terms rather than in minute intensive detail that is characteristic of an advanced course of university calibre. This objective seems to be clearly defined, for no person replying to the questionnaire seemed to be in doubt about it.

2. "To develop ability in the use of the scientific method; e. g.:--
(a) To make accurate observations and to record them systematically.
(b) To draw valid conclusions.
(c) To suspend judgment until sufficient evidence has been obtained.
(d) To develop a critical yet tolerant attitude towards new ideas.
It seems that the first sub-objective is dependent upon two different abilities, that of observing and of recording. The four sub-objectives do not cover all phases of the scientific method. Several persons who replied to the questionnaire made reference to this fact. At least two other sub-objectives should be included in order to round out training in the ability to use the scientific method. These are:

1. To develop the ability to recognize a problem, to see that one exists.

2. To develop the ability to formulate hypotheses from scant data, and its allied ability of developing a theory from more data.
These two sub-objectives are indispensable to the scientific method and should be included under the general objective proposing to develop the ability to use the scientific method. Except for these apparently necessary inclusions the statement of the objective seems to be satisfactory. There does not seem to be any overlapping with the preceding objective although doubtless material from the first objective must be used to develop the ability to think scientifically.

3. "To develop the ability to perform simple experiments and thus to appreciate the experimental basis of science" seems at first glance to belong under the second objective. When we examine the rider "and thus to appreciate the experimental basis of science" we might well wonder if this appreciation is not developed adequately by the second objective. Eleven persons replying to the questionnaire made unsolicited statements of this nature. If the objective were set for purely "mental" development this comment would be quite true and the objective could be included under the second one. The development of ability to use the scientific method is largely a mental development; perhaps it can be sublimed into a purely mental process. The third objective tries to direct us away from this to the developing of the mechanical skills and the concomitant confidence necessary for each of us to do experimenting on our own. It suggests that the student should be able to do experimenting for himself on his own initiative concerning simple problems that arise in his everyday life. The objective does not imply that he must experiment to find out everything, as this would be a very wasteful, tedious, and lengthy education. The student should be able to experiment by himself to the extent that he really can appreciate the contributions of great
scientists. No one can really appreciate these increments to our social wealth and knowledge until he has met the difficulties, doubts, confusion, annoyances, and despairs that go with actual experimenting. The average class-room experiments, wherein the student by himself or in a group does the actual experimenting, is the first step in this direction. He may not learn any more basic principles, nor understand them more clearly than when taught by well-prepared demonstrations but he does obtain training in mechanical skill, and develops self-confidence in various degrees. Here would be trained and tested a pupil’s ability to plan his attack on a problem of not too great difficulty, and his ability to collect the necessary apparatus or materials and arrange them to conduct his experiment or other activity. The committee in including this objective seems to intend something similar to the interpretation outlined here. Good testing technique must determine the "various degrees" of improvements of this type involving the testing of skills as well as "mental" processes.

4. "To enable the student to counteract superstition and to correct erroneous beliefs through the application of scientific principles". It smacks of the scientific crusader’s spirit endeavouring to arouse the student to set his lance against the hoary head of error. Its basic intent is that the scientific knowledge and training developed should not remain passive, or inert, but should be active, dynamic. Well might we ask what we should do under this heading. There seem to be several sub-objectives that would be profitable to record here, namely:— to recognize errors in logic such as some of the simpler syllogistic forms; to train students to trace true cause and effect
to offset the failing that humans show in using mere coincidence as a basis for conclusions or judgments (This is an extremely valuable training in itself and worthy to be included in any educational programme); to develop the attitude of rigorous criticism of what the true scientist usually terms "pseudo-science"; to render the student immune to the misapplied use of the term science, particularly by certain types of advertisements, and to replace a slavish obeisance to the mere words "science" and "scientific" with an upright, honest view of what is truly scientific. With this analysis of the objective completed, the objective does not seem to transgress seriously on the others. It is supplementary to the first two.

5. "To appreciate achievements in the field of science and the contributions of scientists to the modern world". The objective dealing with the appreciation of the contribution of scientists to the modern world and their achievements in general is a straightforward objective with no more implied than in the direct statement. It does not overlap the previous objectives, and is one of the enrichment type of objectives. It deals primarily with the development of attitudes.

6. "To explore the field of science in order to assist the pupil to choose his vocation" does not overlap any other objective. It would suggest that the student from his experience with the subject would learn where his weaknesses and strength lay in order that he may choose more wisely to take, or not, later training in these fields.

With the increasing complexities of modern science professional training is being pushed into the higher grades of even university level so that relatively few vocations based on science are now available to the high school graduate. His graduation and high school training
in science at best is only a pre-requisite for later work. This objective would mean that the science department must be integrated very carefully with the Guidance programme, an admirable objective if carried out properly.

7. "To provide materials for the worthy use of leisure" is an objective at first innocuous in appearance. Commonly, we would suppose that it refers to the development of hobbies, reading interest, and such. The difficulty here lies in the interpretation of the word "worthy". What one society deems worthy another does not; for instance Germany now has her scientists working furiously to develop explosives, to discover all manner of things to render Germany self-sufficient. Why? Not with the view of true self-sufficiency, which is a moderately worthy aim to most people, but for the purpose of gaining European hegemony, if not world domination. Of course most of the other nations in Europe oppose this. Is the great interest which vast numbers of Germans take in science of this type a worthy interest? Will her ability along this line give to her the supremacy which comes from the survival of the fittest, or will it lead to the extinction of a civilization? The worthiness of this aim is difficult to establish, to the subjective nature of worthiness. Take as another example the case of a young girl with the usual biologic urges might conceivably transgress present social mores, possibly abetted by frequent attendance at the lower type of dance halls. Dancing itself is not iniquitous. This is one use of leisure. But suppose as a result of her interests here the girl gave birth to a

---

1. Suggested as a criticism of the objective by Mr. R. Straight, Supervisor of the Bureau of Measurements, Vancouver School Board Office, Vancouver.
child, society at present immediately would condemn her and the incident. In Germany Hitler has announced that persons must take a much broader view of illegitimate birth than they have done in the past and must come to look upon it as the normal procedure. Does not this action of the girl insure that the race will be maintained, and do not the statesmen of the various nations uphold the maintenance of national vigor as a most worthy aim? One society condemns, another accepts the same incident or interest as worthy. Who is to judge worthiness?

Nevertheless, some workable definition of what is worthy would help the teacher even if this definition might not please all philosophers. The suggestion, based frankly on a utilitarian philosophy, is herewith advanced that the development of any interest which does not injure the individual physically, intellectually, and morally (on the accepted standards) and which does not injure other persons would be a worthy interest to follow. There is a wide choice of opportunities in science to develop worthwhile, informative, or harmless interests. To show to the student these opportunities is the main end of this objective which does not seem to encroach on other objectives.

8. "To develop the desire to read scientific literature". It is quite different from the objective "to develop the ability to read scientific literature". The development of desires is very difficult to gauge, whereas the measuring of ability is relatively simple. This objective does not seem to cut across any others, although of course it is dependent upon some of them, notably the first one which develops the understanding necessary for reading.
9. "To develop resourcefulness and adaptability to new conditions" is a very worthy aim. It is based in part upon past experience, and hence the science course could be of inestimable value in giving a wide, sound experience, and in part upon native intelligence which purportedly cannot be developed, much though it may vary in individuals due to biologic functions. It would seem that one of the very important sub-objectives here would be to develop a true appreciation of the understanding of cause and effect, for resourcefulness is not manifest except in new conditions where an analysis of causes is paramount, however rough and ready this analysis might be. A mass of sound information would also be of great value, and this would be contributed by achievements toward the first objective.

10. "To acquire knowledge which will contribute to public and personal health" is an objective which logically must be subsumed under the first objective as it deals primarily with reactions to environment. It merits separate mention from the practical point of view because there exists a great tendency to scotch this aspect in the "straight" science courses. This statement is supported by the results of research conducted by Dr. D. O. Baird who made an analysis of biology books and the courses of study in New York state. One of his findings was that there is less work in pupils' notebooks on health aspects of biology than was found in text-books and courses, suggesting that this work is passed by in some degree. And biology has the greatest amount of time spent on this phase of any of the special sciences. The usual chemistry course ignores health aspects other

than the breathing of pure air, the purification of water (mainly chemical at that), and sometimes mention of drugs that can be synthesized, with some of the newer books including foods and vitamins. Physics passes by nearly all aspects of health excepting the problems of ventilation. In no physics book which has been examined (of the nine leading Physics texts on the continent) is a mention made of care in handling electrical appliances, connections, and such, nor is any mention made of first aid for electrical shocks.

This neglect of health aspects by physics courses and to a slightly less degree by chemistry courses is a flying into the face of research results obtained by Dr. G. S. Craig who found that the weighting given by laymen to an extensive list of objectives of education placed the "acquaintance with such elementary laws of nature as are necessary for the health of the individual and the community" above all others, followed very closely by the "major causes of ill health and the contribution of science to the correction of these causes".

Evaluating the Importance of the Present Objectives

Having examined the objectives of the courses in question, it was next necessary to gain some idea of their relative importance. Admitting at the outset the weaknesses of a general opinion which often expresses the status quo and may exclude very radical, new, and perhaps more valuable ideas, still the best way of obtaining this information of weighting is by a measure expressed by the more enlightened group of human beings. In this group should be placed those who are most affected by the objectives, the teachers of science, and the various persons who are in executive positions.

or administrative ones. Other teachers would form another useful group for purposes of comparison as well as the professors at the University. The laymen if they could be prevailed upon to express opinions would form a sort of control. It was proposed to find this information by the questionnaire method, and to submit the questionnaire to these groups.

(a) University professors; persons who hold high administrative positions. (If enough returns had come in to warrant separation, these two groups would have been segregated.)

(b) Science Revision Committee members; science teachers in the high school. (Again if sufficient returns had come in these two groups would have been separated.)

(c) Other teachers of the high school.

(d) Any other returns not included in above.

Before setting the questionnaire, sections of two books were studied. These provided considerable guidance. The first rough draft of the questionnaire was prepared and submitted to several persons for criticism.

As an outcome of this criticism it was thought advisable to prepare an explanatory sheet to accompany the questionnaire. On it was placed the purpose of the questionnaire, the evidence of its need as indicated by various quotations, and a paragraph of directions to follow in making returns. The objectives were placed on the questionnaire form unaltered and without comments. Only their order or sequence was altered, lest any unconscious influence might be exerted to suggest that the Science Revision Committee had evaluated them in the order in which they appear in the Programme of Studies. The questionnaire was memographed and distributed.

1. Monroe and Englehart, "The Scientific Study of Educational Problems, Chap. 111, Xlll
2. Alexander, Carter; "How to Locate Educational Information and Data
personally, and by mail with suitable postage and addressed return envelopes enclosed. A copy is included in the Appendix.

When the questionnaires were returned they were sorted into the two classifications of Science Teachers and Non-science Teachers for those that were received from the Summer School classes and those that were returned through the mail. The special groups were canvassed individually, so that their returns could be kept separate from the others, and the forms for these were modified slightly in the headings to comply with the positions occupied by those canvassed. The returns from the University professors were not very numerous mainly because most of the science departments are not very active during the Summer Session, and few of the professors were present. Their returns have been incorporated with those who occupied administrative or executive positions. The members of the Science Revision Committee are widely scattered during the holiday season and it has been impossible to secure the opinion of some of them.

As questionnaire returns go they have been rather satisfying for from the one hundred forms issued to date eighty-five have been returned. Of these fifty-seven came from the two Summer School classes, Professor Wood's 1 graduate class in Problems in Education, and Mr. Lord's class in School 2 Law and Administration. The fifty-seven were almost evenly divided between Non-science Teachers (29) and Science Teachers (27) taking as the criterion of segregation the person's one statement as to whether he considered himself a science teacher or not. A few of the Non-science group had taught science but not a great number.

The returns of rankings were totalled, and then averaged; this was thought to be a sufficiently accurate method for the needs. The percentages

\[ \text{Education 23} \]
\[ \text{Law 12} \]
were treated the same way. The ranking returns were further analyzed by making a histogram, then showing the range of variation, and lastly calculating the standard deviation to get an accurate view of the homogeneity of opinion. The results from these two groups follow in the tables.

The following frequency histograms show clearly the massing of opinion in some of tables and a dispersion in others. There seems to be a considerable degree of unanimity of opinion between the science teachers and the non-science teachers, as both groups show almost the same massing or dispersion. The massing of opinion concerning the present objectives 1, 2, and 8 as shown respectively in d, f, and g is very emphatic. Note the peculiar schism in # 3 (j) and 5 (e). In # 4 (b) there seems to be a group who have very emphatic ideas that the overcoming of common errors and superstitions is not the function of a science course. There might be interesting reasons for this but the questionnaire did not call for any comments on the objectives as laid down in the course.
## TABLE 1

Summary of returns submitted to two classes of teachers attending the Summer Session of the University of British Columbia, 1938.

<table>
<thead>
<tr>
<th>Objective</th>
<th>Ranking (Column A)</th>
<th>Percentages (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sum of Ranking</td>
<td>Average Rank</td>
</tr>
<tr>
<td>a Worthy use of leisure (#7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>138</td>
<td>5.11</td>
</tr>
<tr>
<td>Non-science</td>
<td>153</td>
<td>5.46</td>
</tr>
<tr>
<td>b Counteract superstitions (#4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>166</td>
<td>6.15</td>
</tr>
<tr>
<td>Non-science</td>
<td>161</td>
<td>6.46</td>
</tr>
<tr>
<td>c Develop resourcefulness (#9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>135</td>
<td>5.8</td>
</tr>
<tr>
<td>Non-science</td>
<td>127</td>
<td>4.53</td>
</tr>
<tr>
<td>d Body of Scientific know. (#1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>52</td>
<td>1.92</td>
</tr>
<tr>
<td>Non-science</td>
<td>68</td>
<td>2.43</td>
</tr>
<tr>
<td>e Achievements in field of science (#5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>159</td>
<td>5.96</td>
</tr>
<tr>
<td>Non-science</td>
<td>167</td>
<td>5.96</td>
</tr>
<tr>
<td>f Use of scientific method (#2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>68</td>
<td>2.52</td>
</tr>
<tr>
<td>Non-science</td>
<td>72</td>
<td>2.61</td>
</tr>
<tr>
<td>g Desire to read scient.liter. (#6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>216</td>
<td>8.00</td>
</tr>
<tr>
<td>Non-science</td>
<td>256</td>
<td>8.43</td>
</tr>
<tr>
<td>h Public and personal health (#10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>157</td>
<td>5.81</td>
</tr>
<tr>
<td>Non-science</td>
<td>159</td>
<td>4.96</td>
</tr>
<tr>
<td>i Explore re vocations (#6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>197</td>
<td>7.29</td>
</tr>
<tr>
<td>Non-science</td>
<td>185</td>
<td>6.61</td>
</tr>
<tr>
<td>j Ability to experiment (#3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Science T.</td>
<td>189</td>
<td>7.00</td>
</tr>
<tr>
<td>Non-science</td>
<td>189</td>
<td>6.74</td>
</tr>
</tbody>
</table>

Sum of choices made by adding all the rankings for any one objective. Average obtained by dividing sum of choices by frequency; 27 for Science Teachers, and 28 for Non-science Teachers.

Standard deviation calculated by Pearson product-moment method.

Percentages were summated and averaged from returns on questionnaire.
FIGURE 1

Frequency distributions of rankings or evaluations given to the ten objectives of General Science IV and V by Fifty-five Secondary School teachers attending the Summer Session of the University of Brit. Col. (Histograms of "Science Teachers" in red, "Non-science Teachers" in green)

Objective "a" (#7, p159)

Objective "b" (#4 in Bulletin I)

Objective "c" (#9 in Bulletin I)

Objective "d" (#1 in Bulletin I)

Objective "e" (#5 in Bulletin I)

Objective "f" (#2 in Bulletin I)
FIGURE 1 (cont.)

Objective "g"
(#8 in Bulletin I)

Objective "h"
(#10 in Bulletin I)

Objective "i"
(#6 in Bulletin I)

Objective "j"
(#3 in Bulletin I)
### Table 11

Evaluations and Weightings Made by Sixteen Administrators, Principals, and Science Committee Members

<table>
<thead>
<tr>
<th>Objective</th>
<th>Ranking (Column A)</th>
<th>Percentages (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sum of Rankings</td>
<td>Average Rank</td>
</tr>
<tr>
<td>a</td>
<td>Worthy use of leisure (7)</td>
<td>92</td>
</tr>
<tr>
<td>b</td>
<td>Counteract superstitions (4)</td>
<td>108</td>
</tr>
<tr>
<td>c</td>
<td>Develop resourcefulness (9)</td>
<td>75</td>
</tr>
<tr>
<td>d</td>
<td>Body of scientific know. (1)</td>
<td>23</td>
</tr>
<tr>
<td>e</td>
<td>Achievements in field of science (5)</td>
<td>6</td>
</tr>
<tr>
<td>f</td>
<td>Use of scientific method (2)</td>
<td>32</td>
</tr>
<tr>
<td>g</td>
<td>Desire to read scientific literature (8)</td>
<td>120</td>
</tr>
<tr>
<td>h</td>
<td>Public and personal health (10)</td>
<td>88</td>
</tr>
<tr>
<td>i</td>
<td>Explore for vocations (6)</td>
<td>113</td>
</tr>
<tr>
<td>j</td>
<td>Ability to experiment (3)</td>
<td>101</td>
</tr>
</tbody>
</table>
### TABLE 111

**Summary of Evaluations by 78 Teachers**

(Composite of all returns.)

<table>
<thead>
<tr>
<th>Objective</th>
<th>Ranking (Column A)</th>
<th>Percentage (Column B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sum of rankings</td>
<td>Average rank</td>
</tr>
<tr>
<td>a Worthy use (7) of leisure</td>
<td>431</td>
<td>5.53</td>
</tr>
<tr>
<td>b Counteract (4) superstitions</td>
<td>502</td>
<td>6.43</td>
</tr>
<tr>
<td>c Develop resourcefulness</td>
<td>356</td>
<td>4.56</td>
</tr>
<tr>
<td>d Body of scientific knowledge of envir.</td>
<td>155</td>
<td>1.98</td>
</tr>
<tr>
<td>e Achievements (5) in the field of science</td>
<td>463</td>
<td>5.91</td>
</tr>
<tr>
<td>f Use of scientific method</td>
<td>185</td>
<td>2.37</td>
</tr>
<tr>
<td>g Desire to read scient. liter.</td>
<td>629</td>
<td>8.06</td>
</tr>
<tr>
<td>h Public and personal health</td>
<td>429</td>
<td>5.50</td>
</tr>
<tr>
<td>i Explore for vocations</td>
<td>542</td>
<td>6.92</td>
</tr>
<tr>
<td>j Ability to experiment</td>
<td>545</td>
<td>6.98</td>
</tr>
</tbody>
</table>

Numerals in brackets denote order in Bulletin 1 page 159
FIGURE II

Frequency Histograms Showing the Distribution of the Seventy-eight Rankings Given to the Ten Objectives of General Science IV-V.
(Objectives now reorganized in order of importance)

Objective a
(#{1, page 159, Bulletin I})

Objective f
(#{2, page 159, Bulletin I})

Objective e
(#{9, page 159, Bulletin I})

Objective h
(#{10, page 159, Bulletin I})
FIGURE II (cont.)
Frequency Distribution of Seventy-eight Evaluations of the Objectives.

Objective a
(#7, page 159, Bulletin I)

Objective c
(#5, page 159, Bulletin I)

Objective b
(#4, page 159, Bulletin I)

Objective i
(#6, page 159, Bulletin I)

Objective j
(#3, page 159, Bulletin I)

Objective g
(#8, page 159, Bulletin I)
SUMMARY

The Rearranging of The Present Objectives

From Table 111 it will be seen that the order of the objectives arranged with the most important first should be:

1. To acquire a body of knowledge in the field of science which will enable the student to interpret and appreciate his environment.
2. To develop ability in the use of the scientific method; e.g.:--
   (a) To make accurate observations and to record them systematically,
   (b) To draw valid conclusions.
   (c) To suspend judgment until sufficient evidence has been obtained.
   (d) To develop a critical yet tolerant attitude towards new ideas.
3. To develop resourcefulness and adaptability to new conditions.
4. To acquire knowledge which will contribute to public and personal health.
5. To provide materials for a worthy use of leisure.
6. To appreciate achievements in the field of science, and the contributions of scientists to the modern world.
7. To enable the student to counteract superstition and to correct erroneous beliefs through the application of scientific principles.
8. To explore the field of science in order to assist the pupil to choose his vocation.
9. To develop the ability to perform simple experiments, and thus to appreciate the experimental basis of science.
10. To develop the desire to read scientific literature.

This order now provides a guide for the amount of testing material for each objective. The percentage weightings give a more precise guide,
but they should not be used rigidly. A minus variation of one per cent for the lesser important objectives and a greater dispersion of up to a plus ten per cent for the first objective would seem fair.
CHAPTER 11

COMMON FORMS OF TEST ITEMS AND THEIR FUNCTIONS ON TESTS

Scope of Chapter

Having determined the relative importance of the objectives of General Science, with particular reference to the British Columbia courses, General Science IV and V, it is now necessary to investigate tests and types of questions that compose them in order to find which forms of questions will measure in the best manner achievement of these several objectives. In this chapter a survey will be made of the more common types of questions found on both standardized and ordinary tests in the field of science. This will be followed in Chapter IV by a more detailed study of certain standardized tests to see to what degree these tests attempt to measure achievement of the various objectives. In subsequent chapters an attempt is made to adapt and invent question forms to meet the several objectives more adequately.

The Three Major Types of Tests

It may be said of testing that there are three major purposes:

1. To obtain administrative data. In this category should fall the tests used for the allocation of letter-grades. Tests used to segregate the able students from the less able into respective classes have this function. Many Aptitude Tests fall in this grouping, although some like the Seashore Musical Aptitude Test try to obtain an absolute measure rather than a relative one.

In all these tests the essence is to obtain a distribution, to find the relative placement of pupils measured against the others.

2. To ascertain achievement in a given field: A unit test of the factual matter, a general test of ability to use the scientific procedures developed during the course, or a test of the application of scientific principles to health problems would be examples of achievement tests restricted to certain fields. A test of this type should be a comprehensive one sampling all sections of work of at least average importance.

3. To diagnose student difficulties: This test should be quite extensive as it is designed to locate pupil difficulties. When crucial sections of work are being tested, such as a law or principle, several points attacking the problem from different angles should be employed.

The administrative test may be brief and makes much use of sampling. Every item should have a discriminative value differentiating it sharply from the others. It is ordinarily a short test with the minimum number of items necessary to give reliability. It need not be all-embracing as long as there is present a proper sequence of difficulty of the test items. These items should have their discriminative power measured in terms of standard deviations. For instance in preparing the test the examiner may have test items A, B, C, D all of equal discriminative ability but dealing with quite different topics. He need choose only one of these for his administrative test.

Commonly an achievement test may be defined as one designed to express in terms of a single score a pupil's relative achievement in a given field of achievement. Hawkes, Lindquist and Mann assert that achievement tests should be made on a basis of the discriminative power of the items to provide a progression from low standard deviation values to
high ones, it should be comprehensive. With some parts of the definition
the writer is tempted to disagree, mainly on the basis of the word "rela­
tive". The conception of the achievement test as explained above is based
on pragmatic philosophy. While it is admitted that there is continual
change, particularly during the discovery period of science that the world
is in today, nevertheless one can develop measuring instruments that are
less relative and more nearly absolute than is the letter-grading based on
a normal curve of distribution. There is no purpose gained in making an
achievement test rigorously discriminative, for an administrative test
does this. Rather the achievement test should be a "pupils" test" whereby
he can see how much he has or has not learned or developed. It should not
be a mastery test which more nearly approaches the diagnostic test, in func­
tion, but rather it should be of the nature of a more nearly complete or
absolute standard upon which may be indicated the likely high and low
scores. While the ideal test of this nature should cover all points of a
course such a test is impossible in practice. However, the sampling could
be so extensive and intensive that the reliability of the test will be re­
duced only by an insignificant amount from the ideal. Referring to the
illustration in the preceding section if items A, B, C, D test different
topics they must be included in an achievement test. To insist that items
of no discriminative values should be eliminated from achievement tests is
unfair. There may be fifty items of fundamental importance in a course
that all students, having taken the course, could answer. These must be ex­
cluded from present type achievement tests as not possessing discriminative
power. Their inclusion in a test is looked upon in the nature of padding,
and so it would be in an administrative test, but it is not in an achieve­
ment test for it represents the basic or minimum course more truly than the average score taken from the distribution. Achievement testing surely must measure this. To judge a pupil, or a teacher also, upon only the discriminative questions based on a course is not obtaining a true measure. Discriminative tests take off the cream, but then there is a great amount of value in skim milk.

To broaden the conception of achievement tests would seem to be in line with present thoughts on the matter of making reports on students. If the future report is going to be less competitive and more nearly a report on what the student himself is actually doing, it would seem necessary to measure on an absolute standard. The basic achievement which when put into straightforward questions and not garnished or embellished, has little discriminative power could be included in the absolute standard.

Diagnostic tests should be built on a different plan from either of the preceding. They should be extensive, should not be shortened by sampling, and should have several questions bearing on each of the more important or crucial points such as laws or principles, in order to come at the problem from different points of attack. This is necessary because a law or principle is based usually upon two or more factors interacting, and the misunderstanding of any one factor prevents the understanding of the principle. They should be arranged so as to repeat basic items to see if weakness or strength of pupil is consistent.

Research Limited to Questions Testing Achievement

Of these three types of tests only the achievement group is selected for study. It is intended to try to find or devise techniques of testing that may be used in measuring achievements in the fields of the various
objectives rather than to make a statistical or mathematic treatment of
the field, which is another problem entirely.

In all the types of tests mentioned above the forms of questions are
the same. It is not the form of question that makes an achievement or an
administrative test but the plan and construction of the test as a whole.
The form of the question has a considerable bearing upon the validity of
the question and hence indirectly upon the validity of a test. It is im­
perative that the question be analyzed on this basis, "Is it actually
testing what it is desired to test?" If the present form of question does
not test directly what is desired perhaps another form would be better.

Chief Forms of Questions in Common Use

The chief types of questions in common use in class-room work and
upon standardized tests are:—

1. True-false and modifications.
2. Multiple choice and modifications.
3. Completion or Recall, both sentence and paragraph.
4. Matching or Association.
5. Analogy.
7. Catechism.
10. Checking, check lists, identifications.
11. Performance or "Practical" tests.
13. Rearrangements.
15. Constructions, Drawings.

**The True-False Type**

The form of this test causes the student to decide whether a statement is true or false and to register his decision accordingly.

There are many variations of this: Right or Wrong, Yes-No, Plus-Zero, Plus-Minus.

Another type is to arrange a test with the same point repeated in various forms throughout the test in order that consistency of opinion may be checked.

Another form has the special phrase or word tested in italics or underlined. If the student decides that it is false he must correct the phrase or word.

This is one of the first developed objective questions. During the "twenties" of this century it became very popular, but its popularity is waning now. Few standardized tests now employ it. They are rapidly done by students, requiring ten per cent less time than a two-choice question (multiple choice arrangement), up to fifty per cent for recall and for the simpler calculation type of problem not involving complicated computations. Ruch and Stoddart found that the true-false type of question was the least reliable of all types of objective questions, with a reliability for fifty items of only .555 as against .811 for recall.

When this was equalized by the Spearman-Brown formula to give the reliability on the basis of equal time, then the reliability increased slightly to .664 as against .896 for the completion recall types. To obtain a reliability equal to that of fifty recall or fifty multiple-choice questions, another seventy to ninety extra (120-140 total) true-false questions are needed. It can be seen from this investigation that true-false questions take more time to do the same job of measuring as reliably as the recall of multiple-choice do. Töps' work corroborated the first reliability findings but did not agree on the reliability-time basis.

This form of question can be applied to most factual material. It can be of considerable value in testing superstitions, common errors, gullibility, and similar related topics. It is acceptable for testing opinions, and can be utilized for factual tests, particularly where only two possible choices exist.

It is commonly supposed that sets of these questions can be prepared very easily and speedily. When prepared in a hurried manner the quality of the questions is very variable. Questions prepared this way frequently contain familiar phrases and leads, putting too great a premium upon the ability of the student to recall words and not ideas. Tests of true-false types stress texts and words too much for use in science. On the other hand, when sets of questions are prepared carefully by omitting all leads and by qualifying the debatable points sufficiently, there is little, if any, saving of time for the teacher. There is a loss of time if the teacher prepares the extra number of items necessary to obtain reliability equal to the others.

The true-false question often is ambiguous, either too extreme or too expensive in scope, or qualified so much as to be worthless.
Multiple Choice

This type is usually a direct statement followed by several answers of varying degrees of accuracy, one of which usually is correct, or superior to the others. The choices can be short (single words) or long statements. The type is modified sometimes into the multiple response form wherein more than one response is to be identified as correct. Occasionally no response is correct. Almost all standardized tests use this type, some tests employing them to the exclusion of all others. It possibly is the most valuable type because it can be so arranged as to cause active thinking, comparison, reference to data, and reasoned judgments. But merely because a question is multiple choice in form does not mean that it is of the best type. It may be demanding only a recall of facts as in the following:

At sea level air pressure per square inch is (a) 13.6 lb.,
(b) 5 lb., (c) 62.5 lb., (d) 14.7 lb. (e) 30 lb. (f) 29.9 lb.

The above example illustrates how a multiple choice can be made into one that demands the sorting out of a great number of terms or quantities that the student encounters in science. Ruch and Stoddard found that the reliability of this type of test varied considerably according to the number of choices, two-choice questions showing .737, three-choice ones only .598, (four-choice not mentioned), five-choice .796. When corrected for time and reliability the results were in order .902, .806, --, .901 as against .896 for completion recall and .664 for true-false. Toops found that they required sixty per cent more time to do, but gave them a

1. Ruch and Stoddart, op.cit.
time-reliability value of only .607 as against .664 for true-false.

These questions can be applied in testing achievements in almost all fields. They are particularly useful in testing the application of principles, for alternate responses can be made very plausible. The multiple-correct response type is also of good value for testing in science where often several factors are involved, or several consequences may result from a certain cause. It is valuable in testing hypotheses based on data.

To make these questions most valuable great care must be exercised in these ways:

1. Make all answers approximately the same length.
2. Do not become habituated in placement of the correct response.
3. Make the incorrect responses appear quite plausible. Do not make them appear ridiculous so that the correct response stands out from them.
4. Avoid clues to the answer in such apparently neutral things as articles, adjectives, phraseology.
5. Avoid splitting the question by having the responses in the middle. Place responses at the end.
6. For ease in marking have students select the letter of the response and place it in a blank at the right hand side (or left). Avoid underlining the answer in order to speed up marking, and to avoid problems that arise from underlining a part of a response.
7. Make the responses involve active thinking by creating problem situations.
8. Fix the level of difficulty equivalent to the grade or course.

Completion or Recall

The form of this may vary from one blank to be filled in by the necessary answer to complete a sentence, to a paragraph with several to many blanks which must be answered correctly to make a complete correct paragraph. In Chemistry the completion of equations is essentially of this type. Some mathematical problems are of this type when reasoning by analogy is not involved. For example:

To drag a sack of potatoes twenty feet across a floor and using forty pounds of force requires that ———— foot pounds of work be done.

Table completion is another variation of this in very compact form.

This is possibly the oldest form of objective question developed by educational research and one of the most valuable. It is easy to make and need not be bookish. The questions are applicable to every situation which demands a factual answer, to computations, to descriptive questions, to compact tabulations, and to many questions involving the drawing of a conclusion from data presented. This last reference suggests a possibility of its use in testing achievements in the use of the scientific method. With longer responses the objectivity of the answer is usually lessened.

For the short answer type Ruch and Stoddart give a reliability of 1.81 for fifty questions. Toops reports that they take more time in the ratio of twenty-three to twelve (nearly twice as long) than the true-false, When equalized by the Spearman-Brown formula on a time basis Toops reports a reliability of .618 and Ruch and Stoddart .896. The greater amount of time needed to do these questions is due to two factors, primarily because

1. Ruch and Stoddart, op.cit. p.89-103
2. Toops, op.cit., p. 89-52
of the direct recall of the response which sometimes demands considerable mental effort, and secondly the mechanics of writing out a response instead of selecting a letter or a choice.

To be most useful these questions should be prepared carefully.

Arrange the questions so as to:

1. Keep the responses short.
2. Work clearly so that the type of response is shown distinctly.
3. Avoid bookish statements or habitual expressions of the teacher.
4. Include no more in the statement than is functional.
5. Do not confuse the issues. Test for one specific point at a time, e.g.,

   The Danish scientist Oerstad discovered that wire carrying an electric current produced "\[\text{\ldots}\]".

This contains irrelevant material in the reference to Oerstad that may give a lead to the answer. The question as it stands is combining two ideas that should be separated, namely the indirect social contribution made by Oerated and the "straight science." The question reworded to test the science only would be: "A wire carrying an electric current produces \[\text{\ldots}\]." As it stands now it tests science only but it is not precise. The current may produce heat internally; light if the resistance and temperature is high enough, or an electromagnetic field, or all. The question would be more specific in this form "Around the wire carrying it, an electric current produces \[\text{\ldots}\]." This is now precise and eliminates the irrelevance of social contributions.

6. Try to arrange blanks so that they fall at the end of the line and can thus be arranged in columns. Failing this, number the blank and provide a numbered space at the right for the response.
7. Do not leave keys or clues in the composition or grammar of the test item.

8. Because there is a tendency for labelling or naming in this type of question try to word it to make questions more active. Rather give the name or term and have the student apply it, or complete the description.

9. In the tabular forms the questions should be stated separately and a briefer parallel column heading given.

10. In the tabular form avoid simply factual recall and try to include comparisons, reasonings, etc.

11. Do not abbreviate a statement to such an extent that it becomes a puzzle to the student, nor insert so many blanks that he cannot follow the thread of thought. There may be several words of an original statement that may become the question response, but in cases like this choose the word which emphasizes action or understanding of a principle rather than terminology.

Paragraph completion questions are essentially the same as the sentence completion. Great care must be exercised lest the paragraph become too dissected with blanks.

Matching Questions or Associations

One list contains items which must be matched with other items in another list, usually by transposing the letter or number of one to the blank in front of the other. Other variations are: to give the possible responses in a group, from which the correct responses are selected and written in the appropriate blanks.

These questions have been badly prepared in many cases so that the reliability varied. Too many leads were included in many, others lacked
specificity. When prepared carefully they possess a high reliability and objectivity. They are speedily written, and require about the same length of time as the true-false questions take. They can be made quite searching by adding a third or fourth column of information to be matched against the first and second.

They are not very effective when trying to test complex principles, nor for the interpretation of data. They can be used to better advantage when testing brief factual relationships such as matching men's names to their discoveries, laws with their applications, causes and effects, descriptions, conditions or demands with vocations, attitudes desirable in certain situations, and similar material.

The matching questions do not take up as much space as the multiple-choice, or even as much as the true-false usually. They are relatively easy to make and to mark. Observe the following rules when preparing them:

1. The prime rule is to have homogeneity of choices. All responses of one column should be plausible answers for the other.
2. Leads must be eliminated, and stock phrases also.
3. There should be about ten to fifteen items with a minimum excess of three possible responses in order to avoid answering by elimination.
4. One or two responses should be used more than once (with a warning).
5. Arrange items in a manner that reduces mechanical effort to a minimum; names in alphabetical lists; dates in chronological order; long items on the left-hand side, short ones on the right.

**Analogy Questions**

This form is arranged like a continued proportion question, but it usually involves only four terms, one of which is missing and must be provided by the student. This form of question usually involves reasoning by analogy. Many calculation problems are of this type; e.g.,

1. If the mercury in a barometer tube stands at thirty inches when air
pressure is 14.7 pounds per square inch, how high will the mercury be when air pressure is 7.35 pounds per square inch?

These are primarily a form of syllogistic reasoning. They are in the nature of continued proportion where the fourth term or extreme can be found from the others. Arithmetical calculations are of this type. These questions are highly objective, speedily done by the student if he knows the material, and fairly reliable, being about on a par with the four or five-response multiple choice.

They are very suitable when testing laws such as Ohm's Law, Charles' Law, Boyle's Law, gravity pressure in fluids, etc. They are also of great value in comparative work in Biology wherein one organism or organ is compared with another. This type of test item is not as common in the higher grades as in the lower. It might be very advisable to extend its use because of the logical training involved.

Because of the brevity of these there is relatively little in the way of irrelevant material, leads. On this basis they perhaps exceed all other forms in value.

1. They must be prepared with care to see that the analogies are fundamental and not superficial.
2. In order to avoid memoriter response the analogies when applied to calculations must not be of the type done habitually in class.
3. If these questions are prepared afresh each time, the great number of acceptable combinations of factors will enable the presentation of new material and thus increase active thinking.
4. The logic should not be "crossed over". This is easier to check in short statements than in long ones.
5. Include sufficient explanation in the captions or directions, with perhaps an example, to let the student know exactly what he is to do.
6. They cannot be applied well to isolated facts, but are very useful to test the knowledge of the interaction of several factors.
Diagrams

The data are presented in the form of a diagram or picture which the student must interpret. Direct questions may be presented (finding the new condition when a certain factor is changed), recall, or interpretation. Students may be asked to answer by selecting a letter representing a part of the diagram, or by naming, or by inferring consequences.

This is a form of question that has been limited until recent time to Biology and later to General Science. This form of question is being used increasingly now, so that it is not uncommon to see Physics tests and Chemistry tests with diagram questions. They can be made very objective; they can be done speedily or slowly according to the amount of mechanical work attached, and they have a high reliability when carefully prepared. They add a variety to a test and can test outcomes of teaching other than factual absorption. They are very useful in testing interpretation of the data presented by the diagram. They reduce the premium on words so that a student relying only on "word memory" is placed at a disadvantage.

The diagram bears a closer resemblance to the original article than words do. Aristotle recognized this in his statement that, "A picture is worth ten thousand words". They sample other abilities than ability to read.

Graphs of data may be used demanding interpretation.

The particular care needed in preparing diagram tests is:

1. To make the diagram large enough and clear enough for comprehension of all relevant details. A diagram should not be less than about two inches in diameter.

2. Avoid artistic embellishments.

3. Sometimes shading to segregate one area from another is very helpful. The use of several colors is better still, particularly in biological diagrams where one tissue may disappear from one region and appear in another place.

4. Use bold clear-cut lines.

5. Stress the function of parts or the identifying parts that perform given functions, rather than mere naming.

6. Be sure that adequate directions are given.

7. Arrows or guide lines when used should end in the region required. In mimeograph work solid lines are less confusing than broken ones. Sometimes it is preferable to stamp the number or letter of the response right in the diagram and have a parallel numbering or lettering system at the right hand side of the page for the responses. This column of responses facilitates marking.

The writer found by experiment that a positive correlation ranging from .27 to .77 existed between diagram tests and the actual performance of various laboratory activities made as parallel as possible. This is not high enough to warrant the claim that the diagram test measures the same ability as an actual manipulation test does. The Stanford Scientific Aptitude Test scores correlate rather highly with later scientific success in science work done at Stanford, yielding correlations of .77 for Chemistry and .95 for Physics. In this test diagrams are an important factor, in some portions becoming the major factor.

Although diagrams possibly are not testing the same factors exactly as actual manipulation there appear to be many factors common to both methods. Diagrams have this advantage that the data or materials presented to each are constant, whereas in actual manipulation tests it is rather difficult to obtain exact equality of apparatus or supplies among students. In both methods a similarity exists in the selection of data thought to bear upon the problem.

1. See appendix, pp xiv to xxvi.

2. Xyve, Stanford Scientific Aptitude Test, passim in literature accompanying test.
The "Pershing Laboratory Chemistry" Test made a considerable use of diagrams. They merit wider use.

"Catechism" Type

This is perhaps the oldest form of objective questions. It was developed by schools and methods attached to various churches in the past as far back as medieval times. It has a direct question to be answered by only one brief answer. It was directed primarily to textbook learning and tended to be stultifying. It was highly objective as far as it went but it remained too rigid. It was quickly prepared, easily and rapidly answered, but it emphasized rote learning to such an extreme that it lost favor with educators.

There is a survival in the short, direct questions aimed at measuring factual knowledge. These are seldom found on standardized tests.

Essay Type Questions

The form of these may be a direct question, or an imperative sentence to "outline, make, explain, describe," etc. In either case it is expected that the student will give a more or less lengthy reasoned response. It is intended usually to test such abilities as the student's power of organizing material, of exposition, of tracing cause and effect, or of memorizing of long passages or processes.

They are seldom found on standardized tests because of difficulties in marking. For many years preceding the last two decades questions were asked in such forms as to demand essay or paragraph responses. The essay question can be either an excellent or a bad one according to whether it stresses active thinking and planning, or simply rote learning. This
difference is extremely hard to detect on paper for much of the value of an essay question depends on what was taught and what questions the student had answered before the test. It is unwise to teach for the purpose of assisting students to answer certain questions. This, of course, applies equally forcibly to any objective type of question. It is the freshness of approach that helps considerably in making essay responses more dynamic. Re-casting questions that may have been used once as classroom activities increases the mental activity on a test. The turning of the point of the question toward the application of ideas or principles is a very effective way of improving the question. In class procedures it is quite likely that the chemical properties of oxygen may be studied, and questions like; "List the chemical properties of oxygen" assigned for study purposes. To revise the question in the following form introduces a more active principle than rote memory and repetition: "What are the chemical properties of oxygen that make it important?" The latter form is a more life-like form, for it has the element of genuine enquiry in it. It demands a certain ability to organize, and to apply knowledge gained to solve a particular problem. If the work is taught with the outline of the latter question a good deal of the organizing and application of ideas will be done by the teacher, hence reducing the value of the question somewhat.

Many teachers feel that the essay tests do serve a more valuable function than the rigid objectivist would admit. They can be used very effectively in testing appreciation and interpretation. They are very useful in testing the student's ability to organize material, providing that the question matter is fresh, and they serve this purpose better.
than the rigid objective questions do which attempt to cover the same
ground because the objective questions have been organized already by
the examiner. The essay tests force the student to approach problems
as wholes, to organize his attack, and to break down the problem into
its components. These questions force the student to select his data
from a mass of information so that they can be made to measure the
student's ability to sift the grain from the chaff. In the functions
of organizing an attack on the problem, the marshalling of facts, inter­
pretation and appreciation the essay type of test seems superior to the
objective or short answer type.

The chief difficulty connected with essay tests is the marking of
the responses in a reliable manner. To improve the essay test three
major methods are possible. First a table of specifications of the exact
objectives of the question should be prepared. Next the question itself
must be in the best form possible. Lastly a key for marking or evaluat­
ing responses should be prepared.

The first step, that of laying specific objectives for the question,
is of course a procedure that must be adopted for all good questions.
"What knowledge do I need to test?" "What organizing ability is demanded?"
"What interpretations is the student able to make?" These are examples
of questions that the examiner must ask himself at this stage. He must
go further and list the knowledge, outline the organization and problem,
and try to gauge possible interpretations.

Having set up his specifications for the question, the examiner can
improve the question form and function by judging it on the bases follow­
ing,
1. It is axiomatic that the question must be based on the student's experience, and not completely foreign to it.

2. The question should provide a new situation or a new approach to an old one. Questions should not be stereotyped or bookish.

3. The question should demand an organization of material on a problem-solving basis rather on a factual recall or descriptive basis (Unless it is these abilities that are to be tested).

4. Sampling should be more extensive than many essay tests have demanded, if one may judge of the great number of examinations of the essay type which are composed of four, five, or six questions to be done in two or three hours. To answer twenty short questions in an essay manner within the same limit of time extends the sampling to four times the original and thus increases reliability. If it is desired to test ability to handle long problems one of these is about all that should be included on one examination.

5. The question should give a clear idea to the student as to what is wanted. Sometimes the breaking of the question into parts or headings is helpful.

6. The question should put a premium on thought rather than verbiage. Students should know clearly that extended wordy answers are not as valuable as are lucid concise ones. A very fine means of illustrating this is to take an example of a wordy, frothy response and a good succinct one. Ask the class to put down one point for each new idea bearing on the question, then read each chosen response slowly. The class scores are usually all the teaching needed to illustrate the difference between a bad and a good essay response.

7. The actual mechanics of writing should be reduced to a minimum, because it is not writing speed or legibility that is being tested but ability to think in the particular field tested.

8. Avoid using optional or alternative questions.

9. Do not set a value upon a question until the keys are made.

Pairs of questions of fair and superior types are chosen as illustrations:

1. (a) How does water freeze? 
or
    (b) Explain the changes that take place when water cools. 
or
    (c) Describe how water freezes in a lake. 

(d) Why must all the water in a pond be cooled to 4°C before any of the water will freeze?
2. 
   (a) Explain how sulphide ores are smelted.

   (b) How is sulphur removed from certain ores in smelting?

3. 
   (a) What are the properties of sulphur-dioxide that affect living things?

   (b) Why is vegetation injured and human health impaired around smelters releasing sulphur dioxide into the air as a result of smelting sulphide ores?

4. 
   (a) Described the complete process of photosynthesis.

   (b) A white leaf, a green one picked about midnight, and a green one picked at noon of a sunny day all are tested for starch by using iodine. Only the last one contained any. By reference to the processes that occur in leaves account for this difference.

   - The last step is the preparation of keys. Because many examiners in the past have failed to prepare keys, the essay test lost much of its value. It became too subjective. A key based on possible correct responses reduces considerably the subjectivity. How much work is expected, of what quality? Should spelling and compositional difficulties be deducted from the science score, or only indicated to the student? In preparing a key the examiner must guard against two tendencies, that of setting too high a standard that is more in keeping with the examiner's training, and the opposite, that of accepting almost anything that is written.

   - There are at least four plans of evaluating essays and paragraph responses in use at present: These are:

1. Take all the examination papers and read all responses to question one; arrange responses approximately into three, five or seven categories according to merit. If three groups are selected, refine these. If seven, re-read the borderline cases. Having done this marks can now be allocated, top marks for the best paper, and zero for no attempt or an attempt that is short and hopelessly jumbled. This was one of the first methods used in getting away from the subjective method of marking. It is pragmatic in philosophical implication and to this extent is on the same ground as most objective tests. The lack of a key somewhat invalidates results, for a key tends to eliminate subjective reactions, especially if the key has been prepared in conjunction with two or three other examiners who discuss each point.
2. The second method of marking essay questions is to note on a list all
valid or correct responses. Occasionally the examiner will be forced
to accept ideas advanced by the student that he had neither expected
nor accepted previously because a strong valid argument had been ad-
vanced. This list of correct items then forms the marking list
against which all responses are measured by giving a point for each
acceptable item. Maximum marks are almost an impossibility with this
method of evaluation because many students advance points not ad-
vanced by others.

3. The third procedure is a combination of the first two. All the cor-
rect response items are recorded as in the second method. During the
examination of responses they are sorted into a desired number of
groups as in the first method. This involves practically no more
work than the second method, for it takes only a fraction of a second
to place the response in one pile rather than another. The best re-
response as measured by the number of correct items or interpretations
receives full marks if an arbitrary value had been set, and the
others pro rata when compared with the best response. This method
is very much more reliable than the common method of reading an es-
say, then assigning a value to it.

A variation of this procedure was adopted by Dr. W. J. Osburn, now
in the Department of Education, University of Washington, in develop-
ing his "Thought Tests" in history and geography. He took all the
possible correct responses given by a large group of pupils to cer-
tain essay type questions, tabulated them, then prepared a key to use
when administering the tests to new students. The items that were
standardized previously became the yardstick for later tests. An ex-
perienced teacher teaching two to four classes of the same grade and
taking the same subject could prepare a key of this nature that
should be usable for five years or more. It is desirable not to ex-
press scores on a percentage basis, but if they must be, percentages
can be calculated against the key. It is better to take distributions,
or more simply "satisfactory" or "unsatisfactory" directly from the
raw scores.

4. The fourth general procedure is for the examiner to write out the
marking key on the basis of his judgment as to what a perfect res-
ponse, a mediocre one, and a poor one should be. This is a dis-
tinctly profitable method where the question demands interpretation or
attempts to measure depth of discrimination. There is the danger
that the examiner will set his standard too high, more in keeping
with his own training than that of the student's.

Hawkes, Lindquist and Mann suggest a plan that may be taken either
as an alternative or as variation of the above. The examiner prepares a
list of all the items or degrees of interpretation, discrimination, etc.,
which he thinks should be included in a perfect answer, then proceeds to
mark responses in proportion to the number of acceptable items or degrees of interpretation.

These methods are rather dangerous to use by one teacher only. A group of teachers working together gives better results, i.e., less subjective and more reliable. Also it is unwise to put this method in the hands of an examiner who has had no experience with the subject or student level for which the test is designed because the direct contact with the abilities of students does much to improve judgment or qualify decisions.

All of the above four methods are not applicable to any or all essay type responses. The first three are useful for testing arguments, reasonings, cause and effect relationships, applications, and rote learning or memory. The last one is superior for testing organization, the preparation of keys or tables, plan of attack, discrimination, interpretations, and such.

All of these procedures can be improved still further by working out some penalty to be exacted when students include wrong responses. An essay response differs from an objective one (except completion types) to this extent that the student shows his own genuine positive and negative knowledge. On an objective test wrong answers are those provided by the examiner for the student to accept or reject. Normally the student may not have thought of any of these. On an objective test a student may suffer from no more than doubt and yet select a wrong answer that is not his own. On an essay test every wrong point is the student's own, so these really should be deducted from the gross score to obtain the net score. In this matter the essay test has a distinct superiority over the objective type, one that is perhaps worth following further.
The essay tests are not as satisfactory for measuring extent of information as are the objective. The objective type give a much wider sampling in the same time and with less effort. The essay test should attempt more to measure depth of understanding.

**Mathematical Calculations, Computations**

This is a type that is not clearly marked off from the completion type on one side nor from the analogy type on another side. The balancing of equations in Chemistry, the application of mathematical laws and formulae, and longer arithmetical computations come in here. While ratio may be used in these, usually other mathematical processes like the concept of an equation are involved as well.

Nearly all science tests include a few problems of this type. They are objective if simple, but become rather subjective in the very involved problems because examiners do not agree on the methods of solution, the succinctness of work, and similar items.

In science they usually represent the briefest way of representing the functioning of some principle. They are a necessary part of testing, but they are often overdone. The questions involving mathematical responses can be padded quite easily with irrelevant data. It is for this reason that most of the poorer mathematical questions are poor; they attempt to test for two objectives at once, the ability to segregate data, and the ability to apply a law or principle. In testing principles the data should be perfectly straightforward, and should not involve unduly awkward computations in the solution.

In most standardized tests in science, mathematical problems (as distinct from the simple completion type) seldom exceed ten per cent of
the number of items, but often they are weighted to exceed this value on the total test score.

Checking, Check Lists, "Identifications"

Usually in this form of test the student is asked to check or evaluate items or information provided on the basis of certain criteria or according to his own feelings.

These are used frequently in connection with attitude tests, personality tests, and similar work. They can be modified to test factual matter, and are readily adaptable for purposes of comparing a number of items. When used in this way Ruch and Rice refer to them as identifications.

The difference between identifications and the usual check lists is that the student is asked to identify a certain thing as having certain properties or characteristics, and on the check list he is asked to evaluate items according to criteria provided. Whether the student actually uses a check mark or selects a letter is not very important. Both of these question types have the essence of a multiple choice form and function in them. Check lists are of value in trying to discover interests. With care in preparation they can be arranged so as to give a profile view of responses.

Performance or Practical Test

This is an attempt to strike directly at some achievement which involves skills, techniques, and even attitudes and interests. The response to a question of this type is not entirely mental. Such exercises as to type accurately at a certain rate, to swim a certain distance, or to set

1. Ruch, G.M., and Rice, G.A.; Specimen Objective Examinations, p. 19
up certain apparatus would be examples of the simpler types of these questions. In this type it is not so much the question form as the response which is distinctive. It is "testing by doing".

In tests of this type the actual form of question may be direct, or it may be a command to do a certain task. It differs from all the preceding because it involves activities other than thinking and writing. It involves muscular coordination, skill in techniques that have been developed in the course, a plan of attack on the problem, and a certain amount of resourcefulness.

This type of test is used more in universities than in secondary schools. To incorporate this into regular testing programmes would be a wise step because it tests directly achievements in the way of skills and techniques. These cannot be tested properly on paper tests. Possibly one of the reasons these tests have not become common in secondary schools is the time it takes to do them, the need of a large room properly equipped, and the need of sufficient equipment for the tests, (whether these are administered in rotation or not). However, there are many simple performance tests that can be adapted. For instance there is little reason why students should be asked to identify the parts of a flower from a diagram when the actual flower can be given. If it is actually functioning knowledge that is being tested, the examination of the flower is vastly superior.

The "practical tests" that form a part of many science courses in universities are a step in the right direction. They need to be prepared with particular care to see that they are in line with the specifications that ought to be listed. There is some difficulty marking responses. This
difficulty varies with the items tested. Identifying the parts of a flower can be as objective as any paper test item; likewise to identify the parts of a pump; or any machine. Where larger problems involving invention or adaptation are set the evaluation of responses becomes more difficult. For this reason the examiner just beginning to use this form of test should choose simple examples. This provides a better distribution or sampling also. These tests seem to have a use in testing resourcefulness.

Oral Examinations, Interviews

Again it is not so much the form of the question here as the method of responding which characterizes this form. The question may be given orally, as is the usual case, and the response is usually oral. However, the response may be written. It is an individualistic test, a "man to man" test and is inapplicable in general for the average teacher confronted with mass production. Much time is involved and this usually forces a restriction of this method to the urgent minimum.

This is a form of examination which is very transitory in its evidence of activity unless the spoken words are recorded on a dictaphone. For this reason evaluation is very subjective and difficult. It is rather limited in use because it takes much time also. It is very valuable in diagnostic work, and for special purposes. It requires considerable ability on the part of the examiner to make it worth while, an ability to keep up with events and even to anticipate them. If it is a mere oral questioning this might be done better on paper. Its value lies in the interplay of minds, the more personal touch which is often of immeasurable value.
As a means of diagnosis it is very valuable. As an accurate evaluation its merits seem dubious.

Rearrangements

In this type of test the student must arrange a jumbled array of data into the natural sequence, or be able to select stages or steps in a complete sequence.

The essence of this form of question demands the ability of the student to trace cause and effect relationships, trace true chronological sequences, or phase sequences. An example would be to arrange the phases of the moon in correct sequence starting with the last quarter, or to give the stages in the metamorphosis of a butterfly. It is useful in testing knowledge about rhythmic phenomena.

It is necessary to have the question worded so that there is no confusion in the student's mind about the meaning of the factors, nor to give irrelevant clues by the inclusion of time adverbs. The series of events or items should not be too long, certainly not to exceed ten, with five to seven a reliable number.

There is a difficulty in marking these cyclic or sequential forms of questions. If the student omits or advances one of the steps of a series but places all the others in the right order, either advanced or retarded, a difficulty looms. A formula to give credit on the basis of \((n-1)\) where \(n\) stands for the number correct in the sequence may be used in marking.

Comparisons and Contrasts

This form of question tests the student's ability to examine data.
data and find relationships. In this sense it involves a rather active mental process. This process seems to be akin to that of drawing conclusions from data, or forming hypotheses, and would seem to be quite valuable for this purpose. It can be made quite objective when the answer is limited to one word, or one term. Sometimes it is varied by asking the student to eliminate the unrelated fact or item.

e. g. Cross out the word which does not bear a close relation to the others of this group.

(a) Steam chest, manifold intake, carburettor, distributor, pump.
(b) Sepals, petals, root hairs, stamens, pistils.

Example of the first form: In one word or term state the group or article to which each belongs or the function performed.

(a) Steam chest, eccentric, slide valve, connecting rod.
(b) Filament, wheel, element, electromagnet.
(c) Sperms, eggs, spores, gametes, zygote.

Sometimes a student may be given a pair of terms or examples which he must compare to find similarities and differences.

When the responses demanded are long the reliability decreases; when short the reliability is high.

Constructions and drawings

These questions require the student to construct a diagram or drawing from memory. They demand a fairly accurate knowledge of relationships of parts together with a certain manual facility or artistic ability. When the artistic element is reduced to the minimum and only outline or simple diagrams are demanded they serve a very useful function. They are used only moderately in sciences, although biological examinations have in-
cluded considerable questioning of this type. Physics tests in dealing particularly with light and electricity utilize this form frequently; Chemistry tests only occasionally do so.

Comments on Reliability

The reliability of tests is a very specific concept. Reliability of a test is restricted to a certain group writing the test under certain conditions. Change any factor and the reliability changes. A word of caution in interpreting the reliabilities quoted is in order although these quotients do suggest a certain difference in value for forms of questions.

Some of the factors affecting the reliability of tests are; length; wording of items, sequence of items on the test, length of responses demanded (objectivity), directions given to students, form of test items, time.

The reliability of all forms of tests seems to be increased when directions not to guess are given to the students. This improvement is most noticeable in true-false questions.

The foregoing study of the forms of questions was made particularly with the aim to study standardized tests to see to what extent they test the more intangible achievements of students in science courses.

To examine the forms of tests to see if the form played an important part in testing achievements was considered a prerequisite step to the work in the following chapter. The experience gained from this survey

1. Ruch, G.M. and DeGraff, M.H. "Corrections for Change and 'Guess' vs. 'Do Not Guess"; Instructions in Multiple Response Tests"; Journal of Educational Psychology, Volume 17 (1926) pp. 368-375
would assist materially in analyzing standardized tests for the extent to which they measure achievements of the various objectives of General Science IV and V.

From the foregoing study two general conclusions can be drawn.

1. Form plays a limited part in testing achievement. Some questions by virtue of their form cannot test certain mental processes, such as reasoning, planning, drawing, conclusions, while other forms may test these. Factual material can be tested by many forms of questions.

2. Within a given form there may exist a considerable variation in the searching power of a question dealing with a particular field. The degree of vigour possessed by the question in any particular form depends in a large measure upon the training and mental activity of the examiner. If the examiner maintains an attitude of stimulating higher mental activities this will be reflected in the spirit of the question.
In the preceding chapter forms of questions were studied to see to what extent they might be used in measuring all the achievements a student may be expected to make in General Science IV and V. In this chapter standardized tests will be examined to see what types of questions are used, the functions of these forms, their relative importance on each test, and to what extent each test measures the achievements resulting from a striving for the objectives of the science courses in question.

In order to make this survey twenty-nine standardized tests in the sciences together with seven matriculation papers in the sciences which were set by the Department of Education were examined. These tests were chosen from the fields of Chemistry, Physics, Biology, and General Science. They range in type and purpose from achievement tests, aptitude tests, to rating scales and interest tests. For purposes of comparison the matriculation examination papers were included in the survey.

Each test was read carefully to discover the types and the functions of the questions with the hope that one, or a combination of two or three, could be adapted for use in the high school courses in general science. This was done by classifying the test items according to the objective with which they chiefly were concerned. Ten rubrics corresponding to the ten objectives of the course had to be provided. The analysis included also an examination of frequency of a particular form of question and a comparison of this with the total number of points on the test.

Following the analysis tables there is a brief summary of these
points and a few comments. After this part samples of the more original, unusual, or more effective types of questions that possess great promise for testing purposes in General Science IV and V have been abstracted and included in this chapter. In each case an attempt was made to examine the function and potentialities of the example.

In the following analysis tables the objectives of General Science IV and V have been rearranged according to their order of importance as found by the questionnaire. The Roman numerals represent the revised order

1. For many valuable suggestions at this point the writer is indebted to Prof. C. B. Wood.
TABLE IX

ANALYSIS OF STANDARDIZED CHEMISTRY TESTS.

(over)
<table>
<thead>
<tr>
<th>TABLE</th>
<th>CHEMISTRY ACHIEVEMENT TESTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Knowledge, etc.</td>
<td>11 Scientific Method</td>
</tr>
</tbody>
</table>

**HARVARD HIGH SCHOOL CHEMISTRY TEST.** (Gerry) Forms A and B

47 Multiple choice & If mathematical
Completion or Recall calculations in
Usual subject matter. 21 items
subsumed here, then
four problems on test.

**COLUMBIA CHEMISTRY TEST, FORM A**
65 True-or-false 12 Equations are hard
Information items
3 Completion items
Multiple response with
true-false items on
each. Information.
Partly a judgment
5 Math. problems.
Equations Completion

**POWERS GENERAL CHEMISTRY TEST, FORM A AND B; 1924.**
FORM A
Part 1-"5-Mult.choice" Equations are in
part test of drawing
information 55 items
5 Math. calculation
55 points
FORM B similar.

**PERSHING CHEMISTRY LABORATORY TEST. FORM A**
Much use of diagrams Observing and re-
to present data from coning ability is
test by #22 through
to make his observa-
ations and draw con-
clusions, hence could be in next column, except that many items are
purely factual; e.g.
#5,6,7,8,15,16,17-21,
67 out of 69 are
completion type with
a few "choice" mixed in

**MALIN DIAGNOSTIC CHEMISTRY TEST " A"**
Tabular "properties" 5 Math. balancing of
test "good but "rigid" equations
12 points
20 Multiple choice
four items, reasoning 7 problems cal-
Completion or recall
17 items, all inform-

The emphasis directly.
upon lab. tech-
iques would
tend to test for
resourcefulness
#11 detecting
unnecessary
factors seems
usable.

Balancing de-
demand some
resourcefulness

Balancing de-
demand some
resourcefulness
<table>
<thead>
<tr>
<th>HARVARD H.S. CHEM. TEST</th>
<th>COLUMBIA CHEM. TEST</th>
<th>POWERS GEN. CHEM. TEST</th>
<th>PERSHING CHEM. LAB. TEST</th>
<th>MALIN DIAGNOSTIC TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>True-false</td>
<td>True-false</td>
<td>Part 1</td>
<td>Part 1</td>
<td>None</td>
</tr>
<tr>
<td>None</td>
<td>None</td>
<td>3 items</td>
<td>5 items</td>
<td>None</td>
</tr>
<tr>
<td>True-false</td>
<td>True-false</td>
<td>mult.-choice</td>
<td>mult.-choice</td>
<td>None</td>
</tr>
<tr>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

The major part of the test falls into this class. Much use of diagrams in lieu of actual things would give a student a good idea whether or not he were suited for experimental work.
CHEMISTRY TESTS (continued)

(over)
<table>
<thead>
<tr>
<th>TABLE (cont.)</th>
<th>CHEMISTRY TESTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO-OPERATIVE CHEMISTRY TESTS, TEST 1 INFORMATION</td>
<td>\begin{itemize} \item Multiple response questions, 40 questions \item Matching, and variations; 110 items \item Information, theories \item facts, terminology, occurrence, etc. \item Completion 65 points \end{itemize}</td>
</tr>
</tbody>
</table>

| CO-OPERATIVE CHEMISTRY TESTS, TEST C (similar to above) | \begin{itemize} \item Multiple long response questions \item Matching and variations. \item Usual types for factual testing; terminology; application of principles. \end{itemize} |

| CO-OPERATIVE CHEMISTRY TESTS TEST 11 ON USE OF SCIENTIFIC METHOD | \begin{itemize} \item Practically none of this test can be called informational. To do it does not necessitate memorizing many facts before-hand. \item Identification \item Great use of graphs forcing accurate observation, ability to interpret data, checking hypotheses, drawing conclusions from data only. \item Graph or other data given. Selection of the best statements concerning these. \item 8 groups of questions averaging 7 to a group. Identification \item Tables of data likewise. \item Suspending judgment \item Critical attitude. \item 15 questions of multiple choice type but long responses reasons. \end{itemize} |

<table>
<thead>
<tr>
<th>CARNABY AND CARLETON</th>
<th>COMPREHENSIVE MASTERY TESTS IN CHEMISTRY 1935</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usual run of multiple choice, matching, completion or recall; equations. 24 pages of tests averaging about 18 items per page</td>
<td>\begin{itemize} \item Whatever there is in balancing equations. \item Math. calculations. \end{itemize}</td>
</tr>
<tr>
<td>TABLE (cont.)</td>
<td>CHEMISTRY TESTS</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
</tr>
<tr>
<td>V HEALTH</td>
<td>VI SCI. CONTR.</td>
</tr>
<tr>
<td>CO-OPERATIVE CHEMISTRY TESTS.</td>
<td>TEST L</td>
</tr>
<tr>
<td>Two or three 5-6 historic, usual types and Mult. choice commercial processes Matching Mult. choice</td>
<td></td>
</tr>
</tbody>
</table>

CO-OPERATIVE CHEMISTRY TESTS. TEST C.
much as above None None None None None

CO-OPERATIVE CHEMISTRY TESTS. TEST 11 ON USE OF SCIENTIFIC METHOD
Only indirectly by type of matter chosen. None Only indirectly by testing for ability to reason from data and be critical None None None None

CARPENTER AND CARLETON COMPREHENSIVE TEST IN CHEMISTRY 1935
Few items of usual types of usual factual types Mult. choice Matching Recall None None None None
TABLE V

ANALYSIS OF STANDARDIZED PHYSICS TESTS

(over)
<table>
<thead>
<tr>
<th>TABLE PHYSICS TESTS</th>
<th>I KNOWLEDGE</th>
<th>I SCIENT.METHOD</th>
<th>III RESOURCEFUL, IV LEISURE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HURD'S FINAL TEST IN HIGH SCHOOL PHYSICS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62 items total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47 out of 68 factual completion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/14 factual mult.-choice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COLUMBIA RESEARCH BUREAU PHYSICS TEST A and B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>144 items total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All printed, no diagrams</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plus and zero modification of true-false</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>94 information items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COOPERATIVE PHYSICS TESTS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KILZER-KIRBY MATHEMATICS INVENTORY TEST</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10/66 factual information completion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUGHES PHYSICS SCALES, INFORMATION S, THOUGHT R, S</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factual information</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 Direct questions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some of catechism type, some providing data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 mult.-choice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>STEWART-ASHBAUGH PHYSICS TEST (Elect, Sound, Light)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 sets of matching questions total 35 items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 group of 5 questions (matching) based on diagram</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 problems involving mathematical reasoning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TABLE (cont.)</td>
<td>PHYSICS TESTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V HEALTH</td>
<td>VI SCI, CONR.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII SUPERSTI.</td>
<td>VIII EXP.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IX VOCATION</td>
<td>X READING</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HURD'S FINAL PHYSICS TEST**

| None | None | None | None | None | None | None |

**COLUMBIA RESEARCH BUREAU PHYSICS TEST A and B**

| None | None | None | None | None | None | None |

**COOPERATIVE PHYSICS TESTS**

| None | None | None | None | None | None | None |

**KILZER-KIRBY MATHEMATICS INVENTORY TEST**

| None | None | None | None | None | None | None |

**HUGHES PHYSICS SCALES, INFORMATION S, THOUGHT R, S.**

| None | None | None | None | None | None | None |

**STEWART-ASHBAUGH PHYSICS TEST (Elect., Sound, Light)**

| None | None | None | None | None | None | None |
TABLE VI

ANALYSIS OF STANDARDIZED BIOLOGY TESTS

(over)
<table>
<thead>
<tr>
<th>TABLE</th>
<th>BIOLOGY TESTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 KNOWLEDGE</td>
<td>11 SCIENT. METHOD 111 RESOURCEFUL. LV LEISURE</td>
</tr>
<tr>
<td>RUCH-COSSMAN BIOLOGY TEST FORM A &amp; B</td>
<td></td>
</tr>
<tr>
<td>Test 1 Mult.-choice</td>
<td>Test III Making Use of diagrams</td>
</tr>
<tr>
<td>items total</td>
<td>accurate observation of diagram and interpreting diagram.</td>
</tr>
<tr>
<td>Test 11 Mult.-choice</td>
<td>4 points on deductions from data on Mendel’s Laws.</td>
</tr>
<tr>
<td>of statements</td>
<td>15 items.</td>
</tr>
<tr>
<td>15 items, partly</td>
<td>Knowledge, recall.</td>
</tr>
<tr>
<td>reasoning, inform.</td>
<td>Test I Paragraph completion 15-20 items.</td>
</tr>
<tr>
<td>General biology</td>
<td>DETERMINING RELATIONSHIPS</td>
</tr>
<tr>
<td>23 Mult.-choice</td>
<td>Detecting Relationships (Classification types)</td>
</tr>
<tr>
<td>information</td>
<td>None v</td>
</tr>
<tr>
<td>Test I1 Mult.-choice</td>
<td>Fact. or recall</td>
</tr>
<tr>
<td>4 options, 20 items</td>
<td>Test I11 Matching</td>
</tr>
<tr>
<td>Test I11 Matching</td>
<td>20 items</td>
</tr>
<tr>
<td>20 items</td>
<td>Test I1V Mult.-choice</td>
</tr>
<tr>
<td>Test I1V Mult.-choice</td>
<td>on four optional statements, 20 items.</td>
</tr>
<tr>
<td>Information</td>
<td>None v</td>
</tr>
<tr>
<td>94 items total</td>
<td>General biology</td>
</tr>
<tr>
<td>17 recall completion</td>
<td>None v</td>
</tr>
<tr>
<td>16 recognition true</td>
<td>None v</td>
</tr>
<tr>
<td>false</td>
<td>None v</td>
</tr>
<tr>
<td>17 multiple choice</td>
<td>None v</td>
</tr>
<tr>
<td>9 best reason</td>
<td>None v</td>
</tr>
<tr>
<td>35 selecting related</td>
<td>None v</td>
</tr>
<tr>
<td>data</td>
<td>None v</td>
</tr>
<tr>
<td>VAN WAGENEN READING SCALES IN BIOLOGY FORMS A &amp; B 1929</td>
<td></td>
</tr>
<tr>
<td>Informational in</td>
<td>Forces student to draw conclusions indirectly from material of test.</td>
</tr>
<tr>
<td>directly from ma-</td>
<td>None v</td>
</tr>
<tr>
<td>terial of test</td>
<td>None v</td>
</tr>
<tr>
<td>True-false 40</td>
<td>None v</td>
</tr>
<tr>
<td>COOPERATIVE BIOLOGY TEST (1935)</td>
<td></td>
</tr>
<tr>
<td>22 Mult.-choice</td>
<td>None v</td>
</tr>
<tr>
<td>6X3 Diagram</td>
<td>None v</td>
</tr>
<tr>
<td>(18 mult.-response)</td>
<td>None v</td>
</tr>
<tr>
<td>33 Matching Groups</td>
<td>None v</td>
</tr>
<tr>
<td>(99 items)</td>
<td>None v</td>
</tr>
<tr>
<td>TABLE (cont.)</td>
<td>BIOLOGY TESTS</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>VII HEALTH</td>
<td>VI SCI. CONT.</td>
</tr>
<tr>
<td>RUCH-COSSMAN BIOLOGY TEST FORM A &amp; B</td>
<td></td>
</tr>
<tr>
<td>Completion and Completion mult.-choice</td>
<td>mult.-choice</td>
</tr>
<tr>
<td>none direct placed on only indirect actual examination by diagrams. not completely valid</td>
<td></td>
</tr>
</tbody>
</table>

A rather good test although not covering all our objectives with testing techniques.

<table>
<thead>
<tr>
<th>OAKES-POWERS</th>
<th>6/66 items</th>
<th>6 identification or checking type</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>MICHIGAN BOTANY TEST</th>
<th>3 items</th>
<th>Usual types /80</th>
<th>None /80</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>information</td>
<td>entirely wordish</td>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COOPERIDER BIOLOGICAL INFORMATION</th>
<th>20/94 items</th>
<th>Usual types /94</th>
<th>None /94</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>information</td>
<td>entirely wordish</td>
<td>None</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VAN WAGENEN READING SCALES</th>
<th>2/16 chosen</th>
<th>Could be utilized for material to read.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>the same way as for the others by choosing reading material suitable to these aims. But then the test is only indirect while the main aim of this is to test reading ability.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COOPERATIVE BIOLOGY TEST (1935)</th>
<th>9 items /140</th>
<th>4 items /140</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mult-choice</td>
<td>Matching.</td>
<td></td>
</tr>
</tbody>
</table>
TABLE VII

ANALYSES OF GENERAL SCIENCE TESTS

(over)
<table>
<thead>
<tr>
<th>TABLE</th>
<th>GENERAL SCIENCE TESTS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1 KNOWLEDGE</strong></td>
<td><strong>11 SCIENTIFIC METHOD</strong></td>
</tr>
<tr>
<td><strong>RUCH-POPENCE GENERAL SCIENCE TEST</strong></td>
<td>FORMS A &amp; B</td>
</tr>
<tr>
<td>70 items of general information, laws, etc.</td>
<td>Mult.-choice type</td>
</tr>
<tr>
<td>the ten diagrams on part two are combinations of information tests, observation tests, accuracy, and interpretation.</td>
<td>None</td>
</tr>
<tr>
<td>Diagrams and figures to work from</td>
<td>Emphasis not so much on books and words.</td>
</tr>
<tr>
<td><strong>DVORAK GENERAL SCIENCE TEST FORMS</strong></td>
<td>T2</td>
</tr>
<tr>
<td>60 points, all informational</td>
<td>None</td>
</tr>
<tr>
<td>Mult.-choice</td>
<td></td>
</tr>
<tr>
<td><strong>POWERS GENERAL SCIENCE TEST FORMS A &amp; B 1927</strong></td>
<td></td>
</tr>
<tr>
<td>All of 100 items mult.-choice with five options</td>
<td>None</td>
</tr>
<tr>
<td>All informational</td>
<td>None</td>
</tr>
<tr>
<td><strong>UNIT TEST FOR &quot;SCIENCE FOR TODAY&quot;</strong></td>
<td></td>
</tr>
<tr>
<td>14 unit tests for information, of the usual type. True-false with italicized emphasis; mult.choice; matching</td>
<td></td>
</tr>
<tr>
<td>One complete test to be given at end of course on scientific method; data with deductions 12 items; Attitudes 27 &quot;</td>
<td></td>
</tr>
<tr>
<td>Good tests in general. Principles and generalizations with deductions.</td>
<td></td>
</tr>
<tr>
<td><strong>COOPERATIVE GENERAL SCIENCE TEST (Underhill and Powers, 1936, 1937)</strong></td>
<td></td>
</tr>
<tr>
<td>Forms 1936, 1937 have total of 150 items mainly factual</td>
<td>2--3 reasoning</td>
</tr>
<tr>
<td>2--3 reasoning questions, problems</td>
<td>None</td>
</tr>
<tr>
<td>None</td>
<td></td>
</tr>
<tr>
<td>84 items are matching, in small homogeneous groups 53 multiple choice (5) 12 points on diagrams.</td>
<td></td>
</tr>
<tr>
<td>84 items are matching, in small homogeneous groups 53 multiple choice (5) 12 points on diagrams.</td>
<td></td>
</tr>
<tr>
<td>Form N (revised 1937)</td>
<td>2--3 reasoning</td>
</tr>
<tr>
<td>None</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td></td>
</tr>
<tr>
<td>80 items total, all questions multiple (5) choice. diagrams.</td>
<td></td>
</tr>
<tr>
<td>80 items total, all questions multiple (5) choice. diagrams.</td>
<td></td>
</tr>
<tr>
<td>TABLE (cont)</td>
<td>GENERAL SCIENCE TESTS</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>V HEALTH</td>
<td>VI SCI. CONTR. VII SUPERN. VIII EXP. IX VOCATION. X READING</td>
</tr>
</tbody>
</table>

**RUCH-POPENCE GENERAL SCIENCE TEST** FORMS A & B  
5/50 items  3/50 items  3/50  
Usual factual Usual types Usual type of questions give more  
Diagram type none none  
Type of test may Mult.choice emphasis

**Dvorak General Science Test**  
16/60 items  
Mult.choice 2-3/60  1-2/60  None  None  None

**Powers General Science Test** FORMS A & B  
16/100 items  9/100 items  3/100 items None  None  None

**Unit Tests "Science for Today"**  
4/50 on average 1 or 2/50 Attitude  
Age. No new tests includes some items. Promising

**Cooperative General Science Test**  
1936 Part I  
(12/84 items)  3/150 items  1/150  None  None  None

20/150 items as total

1937 Part I  
2/150 items  0/150  None  None  None

18/84 items  28/150 total

Form N, 1937  
9/8 items  3/80 items  1/80  None  None  None
TABLE VIII

ANALYSIS OF APITUDE TESTS AND OTHERS

(over)
<table>
<thead>
<tr>
<th>Table</th>
<th>Attitude and Other Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Knowledge</td>
<td>11 Scien. Method</td>
</tr>
<tr>
<td><strong>Stanford Scientific Attitude Test</strong> (Jyve)</td>
<td>Ex. 1-testing desire to tinker or fix things</td>
</tr>
<tr>
<td>No direct informational test.</td>
<td>Five mult.choice and response.</td>
</tr>
<tr>
<td></td>
<td>Ex. E-Suspending judgment; snap decisions.</td>
</tr>
<tr>
<td></td>
<td>Four mult.choice type</td>
</tr>
<tr>
<td></td>
<td>Ex. J—for reasoning (Bicycle, and gears questions) 8 mult.choice.</td>
</tr>
<tr>
<td></td>
<td>Ex. K—Pick out inconsistencies of phrases in a paragraph. Phrases are numbered</td>
</tr>
<tr>
<td></td>
<td>Consistency of point of view is tested to certain extent 5 checking questions</td>
</tr>
<tr>
<td></td>
<td>Testing observation Ex. F—Optical illusions (ranking type)</td>
</tr>
<tr>
<td></td>
<td>Diagrams Ex. K—Arranging data and selecting cogent data or steps. Good. Five mult.response</td>
</tr>
<tr>
<td></td>
<td>Ex. L—Checking missing lines; ability to formulate systematic plan to attack simple problem. 25 checking items</td>
</tr>
<tr>
<td><strong>Stanford Educational Attitude Test</strong> (Jensen)</td>
<td>None</td>
</tr>
<tr>
<td>Nothing for these rubrics, but may have possibilities when modified to suit the vocational aspects or objectives.</td>
<td>None</td>
</tr>
<tr>
<td><strong>Stanford Interest Report</strong> (Cowdery)</td>
<td>None</td>
</tr>
<tr>
<td><strong>Brewer Vocational Attitude Examination For Boys</strong> (Harvard)</td>
<td>None</td>
</tr>
<tr>
<td><strong>Detroit Mechanical Attitude Test</strong></td>
<td>40 items testing knowledge of articles in trades. 35 mult.choice on uses 40 on functions of parts of various machines</td>
</tr>
<tr>
<td>TABLE (cont.)</td>
<td>APITUDE TESTS AND OTHERS.</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>V HEALTH</td>
<td>VI SCI.CO.</td>
</tr>
<tr>
<td>STANFORD SCIENTIFIC APITUDE TEST</td>
<td>None</td>
</tr>
<tr>
<td>STANFORD EDUCATIONAL APITUDE TEST (for students in Education)</td>
<td></td>
</tr>
<tr>
<td>STANFORD INTEREST REPORT BLANK</td>
<td>None</td>
</tr>
<tr>
<td>BREWER VOCATIONAL INFORMATION</td>
<td>None</td>
</tr>
<tr>
<td>DETROIT MECHANICAL APITUDE TEST</td>
<td>None</td>
</tr>
</tbody>
</table>
### Table IX

**Analyses of Matriculation Papers**

(over)
<table>
<thead>
<tr>
<th>TABLE</th>
<th>B. C. MATRICULATION TEST OR EXAMINATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 KNOWLEDGE</td>
<td>11 SCIENTIFIC METHOD</td>
</tr>
<tr>
<td>PHYSICS 1938</td>
<td></td>
</tr>
<tr>
<td>Completion</td>
<td>Math. problems, if these could be classed here, 46%</td>
</tr>
<tr>
<td>sentence answers</td>
<td>52%</td>
</tr>
<tr>
<td>1937</td>
<td></td>
</tr>
<tr>
<td>Completion, sentence type; 45%</td>
<td>Math. reasoning, etc.</td>
</tr>
<tr>
<td></td>
<td>54%</td>
</tr>
<tr>
<td>CHEMISTRY 1937</td>
<td></td>
</tr>
<tr>
<td>Informational, recall memoriter</td>
<td>Equations ??</td>
</tr>
<tr>
<td>Subjective, descriptive</td>
<td>Math. problems ??</td>
</tr>
<tr>
<td>1938</td>
<td></td>
</tr>
<tr>
<td>Informational, recall memoriter, Subjective descriptive, 85%</td>
<td>Problems and equations 15% approx.</td>
</tr>
<tr>
<td>Shorter questions</td>
<td></td>
</tr>
<tr>
<td>BIOLGY 1937 &amp; 38</td>
<td></td>
</tr>
<tr>
<td>Group 1 Functions</td>
<td>Comparison of two given items to find likenesses and differences; seeking relationships; columnar, brief, semi-objective</td>
</tr>
<tr>
<td>Informational.</td>
<td></td>
</tr>
<tr>
<td>Semi-objective.</td>
<td></td>
</tr>
<tr>
<td>20 points, 40 items</td>
<td></td>
</tr>
<tr>
<td>Group 111</td>
<td></td>
</tr>
<tr>
<td>Modified matching</td>
<td></td>
</tr>
<tr>
<td>Criticising on data pairs</td>
<td></td>
</tr>
<tr>
<td>1936</td>
<td></td>
</tr>
<tr>
<td>10 brief comparison subjective questions, short essay</td>
<td>Comparing type of questions?</td>
</tr>
<tr>
<td>Essay type, planning and memoriter</td>
<td>4 Essay type, planning and memoriter</td>
</tr>
<tr>
<td>20 definitions in groups of five</td>
<td></td>
</tr>
<tr>
<td>10 on structures, functions. Short essays.</td>
<td></td>
</tr>
<tr>
<td>TABLE</td>
<td>MATRICULATION PAPERS</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>V HEALTH</td>
<td>VI SCI.CONT.</td>
</tr>
<tr>
<td>VII SUPERSTI.</td>
<td>VIII EXPER.</td>
</tr>
<tr>
<td>IX VOCATION</td>
<td>X READING</td>
</tr>
<tr>
<td>PHYSICS 1938</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>2-3 items</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td>1937</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>1 item</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td>CHEMISTRY 1937</td>
<td>4-5 points</td>
</tr>
<tr>
<td></td>
<td>3 points</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None directly</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td>1938</td>
<td>None</td>
</tr>
<tr>
<td>BIOLOGY 1937-8</td>
<td>2-4%</td>
</tr>
<tr>
<td></td>
<td>1 point</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None directly</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td>1936</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>5% approx. 1 point</td>
</tr>
<tr>
<td></td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Questions about ex-</td>
</tr>
<tr>
<td></td>
<td>periments reputedly</td>
</tr>
<tr>
<td></td>
<td>done in class</td>
</tr>
</tbody>
</table>
### TABLE X

A COMPARISON OF FREQUENCY OF QUESTION FORM ON TESTS STUDIED.

<table>
<thead>
<tr>
<th>TESTS</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J*</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry</td>
<td>65</td>
<td>220</td>
<td>234</td>
<td>297</td>
<td>76</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>918</td>
</tr>
<tr>
<td>%</td>
<td>7.1</td>
<td>24</td>
<td>25.5</td>
<td>32.2</td>
<td>8.3</td>
<td>2.5</td>
<td>0</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Physics</td>
<td>94</td>
<td>40</td>
<td>19</td>
<td>57</td>
<td>50</td>
<td>73</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>358</td>
</tr>
<tr>
<td>%</td>
<td>26.2</td>
<td>11.2</td>
<td>5.3</td>
<td>15.9</td>
<td>13.9</td>
<td>20.2</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>108</td>
<td>119</td>
<td>161</td>
<td>37</td>
<td>33</td>
<td>4</td>
<td>0</td>
<td>51</td>
<td>12</td>
<td>12</td>
<td>525</td>
</tr>
<tr>
<td>%</td>
<td>20.3</td>
<td>22.6</td>
<td>30.3</td>
<td>7</td>
<td>0.6</td>
<td>0.76</td>
<td>9.7</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Science</td>
<td>0</td>
<td>84</td>
<td>363</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>449</td>
</tr>
<tr>
<td>%</td>
<td>18.7</td>
<td>81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
<td>20</td>
<td>65</td>
<td>65</td>
<td>25</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>185</td>
</tr>
<tr>
<td>%</td>
<td>10.9</td>
<td>35.3</td>
<td>35.3</td>
<td>13.3</td>
<td>8.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>267</td>
<td>483</td>
<td>842</td>
<td>458</td>
<td>184</td>
<td>100</td>
<td>25</td>
<td>64</td>
<td>12</td>
<td>102</td>
<td>2,435</td>
</tr>
<tr>
<td>Percentage</td>
<td>11</td>
<td>19.75</td>
<td>34.5</td>
<td>19.2</td>
<td>7.4</td>
<td>4.1</td>
<td>1.2</td>
<td>2.6</td>
<td>6</td>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>

- A True-False
- B Matching
- C Multiple Choice
- D Completion
- E Identification
- F Computations
- G Direct or Catechism
- H Analogy
- I Re-arrangement
- J* Diagrams (number of Figures only) not included in total.
It may be worth while to draw attention to a few points illustrated in Table IX. General Science tests make a great use of multiple-choice tests, almost to the exclusion of all other kinds. Biology tests use them about one third as much and Physics very little. This disuse of multiple-choice questions in Physics seems odd to the writer for he has had no difficulty in preparing questions of this type on Physics. Many excellent thought-provoking questions can be formed using Physics material. Physics tests seem to rely on true-false and computations mostly; Chemistry upon completion and multiple-choice.

A word of information should be added concerning the true-false results. True-false questions were found on the older tests in the main, and very few were found on tests published within the last three years.

On science tests in general multiple-choice is the most widely used form, with analogies and rearrangements least used. This latter fact seems strange insofar as science is supposed to deal largely with comparisons, and with classifying and organizing knowledge.

It is interesting to compare these results with those of a similar nature obtained in 1929 by Ruch and Rice. These men undertook to tabulate all the frequencies of question forms submitted on the tests entered in the national competition of 1927-1928.

<table>
<thead>
<tr>
<th>FORM OF TEST ITEM</th>
<th>PERCENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Completion</td>
<td>29.71</td>
</tr>
<tr>
<td>2. True-False</td>
<td>24.12</td>
</tr>
<tr>
<td>3. Multiple Choice</td>
<td>16.45</td>
</tr>
<tr>
<td>4. Matching</td>
<td>10.67</td>
</tr>
<tr>
<td>5. Identifications</td>
<td>9.17</td>
</tr>
<tr>
<td>6. Computations</td>
<td>1.77</td>
</tr>
<tr>
<td>7. Re-arrangements</td>
<td>1.77</td>
</tr>
<tr>
<td>8. Analogies</td>
<td></td>
</tr>
</tbody>
</table>

1. Table adapted from Ruch & Rice "Specimen Objective Examinations" p. 9
The following table gives a rough abstract on the applicability of present standardized tests to measuring achievements in General Science objectives.

**TABLE XI**

**COMPARISON OF DEGREE WHICH THE 29 STANDARDIZED TESTS EXAMINED MEASURED OBJECTIVES OTHER THAN KNOWLEDGE.**

<table>
<thead>
<tr>
<th>Questions on Objectives</th>
<th>Chemistry</th>
<th>Physics</th>
<th>Biology</th>
<th>General</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>111 Resourcefulness</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25*</td>
</tr>
<tr>
<td>1IV Leisure</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>V Health</td>
<td>0.33</td>
<td>0</td>
<td>27</td>
<td>10.5</td>
<td>0</td>
</tr>
<tr>
<td>VI Contributions of Scientists</td>
<td>1.55</td>
<td>0.5</td>
<td>4.25</td>
<td>2.8</td>
<td>0</td>
</tr>
<tr>
<td>VII Superstitions Errors, etc.</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>VIII Experimentation</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>IX Vocations etc.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>X Reading</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

* This score is attributed almost entirely to the Stanford Scientific Aptitude test.

Van Wagenen Reading Scales were not included in any of these because none could be obtained for Chemistry, Physics, and General Science. The inclusion of their scores would increase the reading percentages by approximately five per cent.
Were it not for the Cooperative Test Service and the Stanford Scientific Aptitude tests on the ability to use the scientific method there would be only a minute amount of testing done on Objective 11 (the ability to use the scientific method).

Nearly all tests are informational to approximately ninety per cent of the value of the test.

When the values in the table above are compared with the values obtained from the questionnaire it will be seen that testing in science should be re-directed rather drastically. Note, however, that the values for "Health" coincide.

It would appear from the foregoing analyses that no present standardized test would serve satisfactorily the demands of a good testing programme in science, nor could a combination of two or three of present tests suffice although some of the uncommon types could be made very useful.

There is no standardized achievement test satisfactory for General Science IV and V. The Stanford Scientific Aptitude test is good but may be too difficult. Something more than the Van Wagenen Reading Scales is necessary for reading. No tests are made to measure ability to use leisure wisely, attitude, nor for vocations, and few attempt to measure resourcefulness and laboratory practice (Pershing). The Cooperative Test Service attempts to make tests to measure the ability to think scientifically.

In view of the emphasis on the objectives of the course established by the returns from the questionnaire it would seem unjust to judge the progress of any student solely on the knowledge or informational basis, or to judge a teacher's ability entirely on the basis of student success in acquiring information and knowledge. It seems that the only way out
of these difficulties is for the teacher or the examining authorities to
devise special tests which will measure achievements and growth in fields
other than the purely informational.

The least satisfactory science tests appear to be the Physics, for
they are too rigid, unadaptable, and narrow in compass. The more modern
Chemistry and Biology seem to be more progressive in form.

In many cases the Biology and General Science tests are very similar
for each makes much use of multiple-choice and diagrams, and is broader
in spirit, for they stress human values more. It is to be hoped that
tests in General Science IV and V will follow the more recent trends in
Chemistry, Biology, and General Science such as those prepared by the
Cooperative Test Service, and the Stanford Scientific Aptitude Test and
the Pershing Test.

Altogether, it appears that present standardized tests have fallen
very far behind the needs of science courses. They stress far too much
the informational aspect and do not stress enough active thinking in
handling new situations in which the student must use his training in
science. Tests could be fashioned after the Stanford Scientific Ap­
titude Test or the Cooperative Chemistry Test to provide this desired
mental activity.

Likewise consideration must be given to measuring growth towards
the other objectives.

Certain types of questions have appeared as the "stock-in-trade"
of examiners. Some few seemed to possess unique functions or at least
are adaptable to measure some of the achievements frequently passed by.
Three sources of test items were most remunerative of effort. These
were the Kuch-Popence General Science Test, the Cooperative series, and
the Stanford Scientific Aptitude Test.

The following pages contain abstracted test items chosen mainly from the above-mentioned three sources in order to demonstrate some of the types of questions that should be useful in a testing programme in General Science IV and V.

SPECIMEN TEST QUESTIONS

From: Ruch-Papenoe General Science Tests, Forms A and B

Purpose of test items: not stated by authors; (mainly for information).

Directions: Fill in each blank so as to make a true statement.

Figure 2.
In the diagram of a typical flower:

a. The petals (the corolla) are marked by the letter...

b. The stamens are marked by the letter ....

c. The sepals (calyx) by the letter ....

d. The pistil is marked by the letter ....

(This test shows a good restraint in the use of shading. It is clear, accurate, and not too idealized.)

Figure 4.
a. In this lever the force is applied at ....

b. The fulcrum is placed at the point ....

c. The mechanical advantage of a lever of this class is always............. than 1.

Figure 10.

a. The mechanical advantage of this pulley system is ........

b. The rule for the mechanical advantage of any pulley system is that the mechanical advantage is equal to the number of times the cord passes to and from the.................. pulley.

c. Disregarding friction, the force needed to lift the 100-pound weight shown is ............. pounds.
Figure 15. (On the test)

A. a. This is a drawing of a ................. ............... 
   b. The piston is lettered ...................... 
   c. The valve which opens on the upstroke is lettered
   d. The greatest distance that valve C can be placed above the level of the water at D, if the instrument is to work successfully, is about ............ feet.

Figure VI

Figure VIIa. A green under-water plant has just been placed in the apparatus shown in the sunlight. The gas which is being collected at A is ...................... 
   b. A glowing splint placed in this gas will ...................... 
   c. This illustrates a phase of the physiological process known as ...................... 

Figure 15, Form B.

a. In the flask shown at A there is a solution of molasses to which some yeast has been added. The gas which is being collected at B is ...................... 
   b. If this gas is passed into ...................... 
      a ................................ will be formed. 
   c. There will be formed in the liquid in the flask marked A a quantity of ...................... 

Applicability: This form of test would seem to be useful not only for informational testing but for testing powers of observation. Possibly it could be used in testing for "simple experimentation", measuring the student's ability to recall what has been done. It is possible to memorize drawings as information can be, but a few new twists to the diagram and new forms of the same old problem will usually eliminate the memorizers. This type of question seems promising for testing the parts and functions of more complex machinery than shown here, and is much used for biological forms.
SPECIMEN TEST QUESTIONS.

From: Pershing Laboratory Chemistry Test, Form A.

Purpose of test items:
To measure achievement of pupils in laboratory technique; to recognize suitable apparatus; to detect errors of procedure, in apparatus set-up.

Directions: Fill in the blanks with the correct answers. Use diagrams for reference.

Metals sometimes may be identified by fusing the metal salt with borax to a transparent bead. Some metals yield a bead of a given color when fused in the oxidizing flame and a bead of a different color in the reducing flame. Thus a compound of iron fused with borax in the flame at "A" of Fig. 6 will produce a colored bead while if fused at "B" will produce a color. A borax bead containing manganese and fused in the oxidizing flame will produce a color.

11. Study the apparatus shown in Fig. 16 and recall characteristics of substances which may be prepared in this type of set-up. Of the substances listed below mark with an "A" those for which the apparatus is "Applicable", and with "N" for those for which the apparatus is "Not applicable".

60. Carbon dioxide ..........63. Nitrous oxide ........
61. Hydrochloric acid ..........64. Hydrogen ...........
62. Oxygen ................

111. A student working in the laboratory desired to prepare and collect carbon dioxide. He decided to collect it by water displacement method. The apparatus is shown in Figure 12. After the reaction had been going on for some time, he failed to collect any gas in bottle B. Check the apparatus and note any difficulties. The student did not collect the gas because: (Check correct answer.)

a. The gas dissolves in water
b. Heat should be applied to the apparatus.
c. Tube F does not extend down into the liquid.
d. Tube D does not extend down into the liquid.
e. Not enough marble in the bottle.
f. More zinc is needed.
g. Acid used was too strong.

(These figures tend to be too small. Compare with preceding.)

Applicability: This type of question seems on the surface to test fairly well for the common techniques, and on the surface should approximate scores with the tests made with the actual materials. The investigator has not been able to find any control experiments on this basis, nor correlation experiments, done by these people. The test seems to fit in with "development of the ability to perform experiments". This problem was the basis
of some research into correlation the results of which are given in a later chapter. The correlations are not sufficiently high to assume that this type of test is equivalent to a "practical test".

**SPECIMEN TEST QUESTIONS.**

From: The Cooperative Chemistry Test, Test II, page 7.

**Purpose of test item:**
To measure the ability of students' interpretation of experimental data or results.

**Directions:** Each exercise consists of two parts: first a description of an experiment and the results obtained; and secondly of conclusions and interpretations of the experiment. You are to judge the soundness of these interpretations. Read carefully and assume that all facts given you are correct. If it is a sound conclusion based on this experiment alone place a (1) in the first parenthesis after each interpretation.

If it is unsound because it is contradicted by this experiment place a (3)

If it goes beyond the results of this experiment only, place a (+) in the second parenthesis (2).

That is, the interpretation may be true or false, but you are unable to say from this experiment only. Next, consider all these interpretations marked two (2) that go beyond the reach of the experiment and further refine your opinion of these in this way:

If you judge the interpretation to be true or highly probably (judging from evidence which you have obtained from other sources than this experiment.) mark it Plus in the second parenthesis (2) (+)

If you judge interpretation to be untrue or highly improbable (judging from sources of evidence other than this experiment) mark it zero (0) in the second parenthesis (2) (0)

If you cannot decide whether the interpretation is probably true or untrue do not place any mark in the second parenthesis. (2) ( )

Plus and zero are used only after interpretation that you mark (2)

(1) ( ) SOUND CONCLUSION; based on this experiment only.
(3) ( ) CONTRADICTED; by results of this experiment only.
(2) (+) HIGHLY PROBABLY; but goes beyond results of this experiment.
(2) (-) HIGHLY IMPROBABLE; goes beyond the results of this experiment.
(2) ( ) CANNOT DECIDE; goes beyond the results of this experiment but cannot decide whether highly probable or highly improbable.

6. The following experiment was performed individually by 160 students. The same amount of copper was heated with an excess of sulphur, forming cuprous sulfide.

Each student weighed his product and found the percent of copper that it contained. The results for the 160 students are plotted in the graph below.

The theoretical percent of copper in cuprous sulphide is 79.9. This percent is calculated from the formula weight of cuprous sulfide.

(Graph and questions on following page.)
8. (continued)

![Graph](image)

**Figure X11**

a. The excess sulfur was completely burned in each of the experiments.

b. Some of the students did not heat the copper and sulphur long enough in the presence of air.

c. When the students performed this experiment, each student obtained the theoretical percent of copper in cuprous sulphide.

d. A constant error was introduced in some of these experiments.

e. More students obtained results below the theoretical percent of copper in cuprous sulphide than above it.

f. On the average the percent of copper in cuprous sulphide found in these experiments was higher than the percentage calculated from its formula weights.

g. Some of the students weighed the cuprous sulphide while it was hot.

h. Students who found that their cuprous sulphide contained 72% copper had more sulphur in their cuprous sulphide than students who found 86% copper.

i. Students who found their cuprous sulphide contained 72% copper had more copper in their cuprous sulphide than students who found 86% copper.

---

**Applicability:** This seems to be a very fine type of question to measure ability to draw valid conclusions from data presented. While it may seem to be unduly lengthy in its directions here, in the real test the directions do not occupy such a proportionate amount of space because seven other questions of equal length are included under the same set of directions.
From: Standard Scientific Aptitude Test.

Objective: To test student's ability to proceed with caution and to read instructions carefully, and his thoroughness of execution. (Authors claim that they really have very little to do with illusions, which at first glance they seem to be.)

11. Rank the rectangles A, B, C, D, E, F, G in order of their height; that is, write 1 in the small space next to the letter in the column corresponding to the highest rectangle, 2 next to the letter corresponding to the next highest, etc.

![Figure XIII]

Figurc XHl

Applicability: Questions of this type can be used to test students' ability to make accurate observations and to record them. The inclusion of optical illusions catches the careless worker. Although this test item might appear to be too easy for many persons' attention can be drawn to the fact that it is used in Stanford University with some success, and to the fact that the question has a good discriminative value. This is a paper test that appears to replace actual measurements but in reality demands that measurements be made. Students have rulers and are permitted to use them but many do not do so taking the question as too easy to warrant such an outlay of effort.
Exercise 0

Objective of test item: To test ability to detect fallacies and not to be misled by apparent plausibilities. (Authors claim that this item is a good "bait" for the imaginative minded who are not scientific, as this particular group does very poorly on these items).

11. At a recent meeting of the American Association of Mechanical Engineers the following project received thorough consideration. With the future development of extremely light gas engines it will be possible to build dirigible balloons much lighter than those built today. It might be then possible, by the use of air tanks, provided for breathing, to attempt a flight to the moon. Supposing that the distance to the moon is 200,000 miles, and the average velocity of such a dirigible would be 100 miles per hour, it would be possible to complete the journey in about 2,000 hours.

The following reasons either for or against the project were given by various members. Put an X in the squares next to the reason which you would endorse and a --- next to those to which you would object.

( ) 1. Less than 25 years ago almost everyone believed that flying, as we have it today, was a rank impossibility. Therefore, the above project is worth trying.

( ) 2. The above project is worthless, for it is well known that air in the upper layers of the atmosphere does not contain oxygen and therefore is not suitable for breathing.

( ) 3. The above project is worthless, for it is definitely known that the atmosphere does not extend beyond a few hundred miles from the earth.

( ) 4. The above project is worth trying, for the advance in engineering is more rapid than ever and it is unwise to set any definite limit to it.

(The manual and scoring key places a minus sign in front of #1, 2, 4, with plus in front of #3 to get full credit of 5; 2, 3, as plus given credit of 3.)
From: Stanford Scientific Aptitude Test, Exercise K

Objectives of test item:
To detect the aptitude of the individual in visualizing a statistical or experimental situation; capacity for analytical discrimination of values of data; of recognizing a likely plan of attack.

Directions
11. A housewife uses 2 quarts of boiling water for her coffee and wishes to find which of her three kettles, a 4-quart aluminum, a 5-quart copper, and a 3-quart granite one, consumes the least gas for the boiling of the water. Check only those statements which will enable her to get the right answer.

( ) 1. Fill all three kettles with water.
( ) 2. Pour into each kettle 2 quarts of water.
( ) 3. Place all three kettles on the gas range, heat them at the same time, and time each kettle until the water begins to boil.
( ) 4. Place all three kettles on the three different burners on the gas range and heat them one after another.
( ) 5. Place one kettle at a time on the same burner and heat it.
( ) 6. Time each kettle until the water begins to boil.

1. A Physicist wanted to measure the length of a fine wire with precision; for this reason he measured it several times. Below are given the results of the measuring:-

1st measure ................. 13.63 cm.

2nd measure ................. 15.13 "

3rd measure ................. 13.12 "

4th measure ................. 13.14 "

5th measure ................. 13.15 "

6th measure ................. 13.16 "

What is the probable length of the wire? Answer here ( )

Applicability: Questions of this type could be utilized in testing a student's ability to select data, to plan the essential steps of a bit of experimenting. It could be used also to test his powers of organizing an attack on a problem under "the ability to perform simple experiments".
From: Stanford Scientific Aptitude Test, Exercise E

Purposes of test items:
To test "suspended judgment versus snap decisions"; to test tendency to guess or act on insufficient data.

Directions:
Place a check mark (X) in the space next to the correct answer below:

1. What will be the average cost of living in this country in the year 3000?
   ( ) 1. About $50 per month per capita. ( ) 4. About $300 per month per capita.
   ( ) 2. About $100 per month per capita. ( ) 5. About $500 per month per capita.
   ( ) 3. About $200 per month per capita. ( ) 6. If unable to answer put a check mark in front of this.

11. If you stack nickels in one pile 10 feet high, it will contain;
   ( ) 1. About $100. ( ) 3. About $225.
   5. If unable to answer put a check here ( ).

111. A certain government, selling land, offered it on the following terms;
   ( ) 1. If the buyer is an immigrant, he may pay $1,000 every year for 20 years.
   ( ) 2. If the buyer is a native born, he may pay
      $100 the first year
      $300 the second year
      $500 the third year, and so on, the annual payment being increased each year by $200 for 20 years.
   ( ) 3. If the buyer is a war veteran, he may pay
      $1 the first year
      $2 the second year
      $4 the third year, and so on, the annual payment being doubled for each year for 16 years.
   ( ) 4. If unable to answer place a check mark here ( ).

Which terms are the most advantageous? Put a check mark in the corresponding square.

Applicability: Questions of this type could be used quite easily to test a pupil's ability to suspend judgment when confronted with insufficient data. This is part of general ability in use of the scientific method.
From: Stanford Scientific Aptitude Test

Purpose of test item: To test student's ability to detect inconsistencies.

Read the following directions carefully. Read the five paragraphs. If a paragraph is consistent throughout, put an X in the square across from the top of each paragraph; if it is not, place a there, and write in the spaces across from the lower part of each paragraph the numbers corresponding to the phrases or sentences which cause the inconsistency or lead to an illogical conclusion.

1. At sea level, when atmospheric pressure is normal
   water boils at 212°F. When atmospheric pressure drops
   below normal, water boils at a temperature lower than 212°F.
   In localities situated above the sea level, atmospheric pres-
   sure is often below normal. In such localities
   water always boils at temperatures below 212°F.

11. When a body is heavier than its volume of water, it sinks;
   otherwise it floats. Cork is lighter than water;
   therefore it floats. Sodium is lighter than water.
   Sodium is a metal. Metals usually sink in water.

A chunk of metallic sodium thrown in water will float.

Applicability: This form of test could serve a very valuable function in measuring ability to think clearly enough to counteract superstition and to correct erroneous beliefs. Erroneous beliefs cling on usually because they are plausible, but they usually contain consistencies which this detects.
S P E C I M E N  Q U E S T I O N S

F r o m :  S t a n f o r d  S c i e n t i f i c  A p t i t u d e  T e s t ,  E x e r c i s e  I

P u r p o s e s  o f  t e s t  i t e m s :
T h e  t e s t  o f  t h e  t r a i t  o f  a " b e n t "  f o r  e x p e r i m e n t a t i o n  h a s  b e e n  d e v i s e d  t o
d e t e c t ,  n o t  t h e  a c t u a l  e x p e r i m e n t a l  a b i l i t y  d e b u t  t o  t r a i n i n g ,  b u t  t h e  f i r s t
i m p u l s e  w h i c h  i s  u s u a l l y  s y m p t o m a t i c  o f  a n  e x p e r i m e n t a l  b e n t .

D i r e c t i o n s
S u p p o s e  t h a t  y o u  h a v e  p l e n t y  o f  l e i s u r e  a n d  t h e  n e c e s s a r y  m e a n s  f o r  t h e
m e e t i n g  o f  t h e  s i t u a t i o n s  d e s c r i b e d  b e l o w .  C h e c k  ( X )  f r a n k l y  t h e
s t a t e m e n t  w h i c h  c o m e s  n e a r e s t  t o  t h e  w a y  i n  w h i c h  y o u r  f i r s t  i m p u l s e
w o u l d  l e a d  y o u  t o  h a n d l e  t h e  m a t t e r .  ( I f  y o u  w i s h  t o  b e  h e l p e d  b y  t h i s
t e s t  y o u  m u s t  b e  a b s o l u t e l y  f r a n k . )

I .  S u p p o s e  t h a t  y o u r  a l a r m  c l o c k  s u d d e n l y  s t o p p e d  b e c a u s e  o f  s o m e
t r o u b l e ,
( ) 1 .  T r y  t o  d e t e r m i n e  h o w  s e r i o u s  t h e  t r o u b l e  i s ,  a n d  t h e n  t a k e  i t
t o  a  w a t c h m a k e r .
( ) 2 .  I n s t e a d  o f  t a m p e r i n g  w i t h  t h e  c l o c k  a n d  m a k i n g  m a t t e r s  w o r s e ,
t a k e  i t  t o  t h e  w a t c h m a k e r .
( ) 3 .  L o c a t e  t h e  c a u s e  o f  t h e  t r o u b l e  a n d  t r y  t o  c o r r e c t  i t .

I I I .  Y o u  w i s h  t o  k n o w  w h e t h e r  t h e  a s s e r t i o n  t h a t  t h e r e  a r e  s p o t s  o n  t h e
s u n ' s  s u r f a c e  i s  c o r r e c t .
( ) 1 .  L o o k  u p  t h e  m a t t e r  i n  a  t e x t b o o k  o n  A s t r o n o m y .
( ) 2 .  A s k  a  c o m p e t e n t  p e r s o n  t o  g i v e  y o u  t h e  i n f o r m a t i o n  d e s i r e d .
( ) 3 .  O b s e r v e  t h e  s u n  t h r o u g h  a  t e l e s c o p e .

V .  S u p p o s e  t h a t  y o u  a r e  v e r y  m u c h  i n t e r e s t e d  i n  t h e  b e h a v i o u r  o f
m e t a l l i c  p o t a s s i u m  i n  w a t e r .  T o  g e t  t h e  i n f o r m a t i o n  d e s i r e d ;
( ) 1 .  L o o k  i t  u p  i n  t h e  E n c y c l o p e d i a  B r i t a n n i c a  u n d e r  " p o t a s s i u m . "
( ) 2 .  L o o k  i t  u p  i n  a  g o o d  c h e m i s t r y  b o o k .
( ) 3 .  D r o p  a  p i e c e  o f  m e t a l l i c  p o t a s s i u m  i n t o  w a t e r .
( ) 4 .  A s k  a  c o m p e t e n t  p e r s o n  t o  g i v e  y o u  t h e  i n f o r m a t i o n  d e s i r e d .

A p p l i c a b i l i t y :  T h i s  q u e s t i o n  t y p e  c o u l d  b e  u s e d  t o  t e s t  f o r  v o c a t i o n a l
l e a n i n g s .  B y  c h o o s i n g  i t e m s  f r o m  t h e  v a r i o u s  f i e l d s  o f  t h e  G e n e r a l  S c i e n c e
c o u r s e s  s o m e  m e a s u r e  m i g h t  b e  g a i n e d  a s  t o  t h e  a m o u n t  t h e  s t u d e n t  h a s  h a d
h i s  l a t e n t  p o s s i b i l i t i e s  a w a k e d .  T h i s  t y p e  o f  t e s t  w o u l d  n o t  s u f f i c e  b y
i t s e l f  b u t  m u s t  b e  s u p p l e m e n t e d  b y  c a s e  s t u d i e s  a n d  j o b  a n a l y s e s .
SPECIMEN TEST ITEMS

From: Cooprider Information Test in Biology

Purpose:
To test for biological information only.

Completion tests or recall:
2. The gas given off by animals in respiration is ....................
16. Thallophytes that have no green coloring matter are known as ...........

Recognition (True-false)

Place a check ( ) BEFORE the sentences below that are true and a cross, ( X ) BEFORE those that are not true:
1. The pistil and stamens are the most important parts of a flower.
15. The tomato is a berry.

Multiple Choice (Underline the best answer.)

1. Rubber is obtained from (animals, oil, minerals, coal).
2. Starch is made by plants in the (roots, soil, leaves, flowers, bark).

Best Reason Modification of Multiple Choice. (Check best reason with )

11. A frog lives in the mud at the bottom of a pond all winter so that:
   1. It will not be seen.
   2. It can reproduce.
   3. It can keep warm.
   4. It will not freeze.

Classifying or Tracing Relationships

In each group of words below draw a line through one word that does not belong there.

2. Scales, endoskeleton, exoskeleton, hairs, feathers.
3. Eye, antennae, nose, hair, tongue.
4. Turtle, alligator, frog, chameleon, lizard.

Relation or Logical Selection Type.

In each group of words below draw a line through two words in the parentheses that tell what the thing always has.

2. Bird (nest, eggs, bones, song, tail)
3. Cell (cell-wall, protoplasm, nucleus, cilia, centrosome)

Applicability. The usual run of tests for information. Quite usable thus.
**SPECIMEN TEST ITEMS**

**From:** Cooperative Chemistry Test, Test II

**Purpose of Item:** To test for the ability to apply principles.

**Directions:** In each of the following exercises a problem is given. Below each problem are two lists of statements. The first list contains statements which can be used to answer the problem. Place a plus sign ( ) in the parentheses after the statements which tell what will probably happen.

The second list contains statements which can be used to explain the right answers. Place a plus sign ( ) in the parentheses after the statements which give the reasons for the right answers.

1. Chlorine is a poisonous gas. A few breaths of air containing as little as one-fifth chlorine gas is fatal. Magnesium chloride is more than three-fifths chlorine. What would happen as the result of eating some magnesium chloride? Explain.

   a. Death will result .... a. ( )
   b. The magnesium chloride might improve the flavor of the food.... b. ( )
   c. There will be no injurious effects .... c. ( )
   d. The chlorine from the magnesium chloride will irritate the lining of the stomach .... d. ( )

From the following statements select and check the ones which indicate the line of reasoning that you followed in making your predictions above.

   e. Elements lose their characteristics when they form a chemical compound. .......... e. ( )
   f. Magnesium chloride is a mixture of magnesium and chlorine.... f. ( )
   g. Chlorine reacts with complex organic compounds. .......... g. ( )
   h. Magnesium chloride is a salt similar to sodium chloride ......... h. ( )
   i. The properties of a compound are largely influenced by the properties of the elements forming the compounds. .......... i. ( )
   j. Small amounts of chlorine irritate living cells and prevent them from functioning properly. .......... j. ( )
   k. The ingredients of a mixture retain their individual properties after being mixed. .......... k. ( )
   l. Chlorine reacts chemically with living cells producing death.... l. ( )

**Applicability.** When modified it should work equally well in Gen. Sc. IV & V, to be used to test development of ability to use the scientific method.
SPECIMEN TEST QUESTIONS.

From: Van Wagenen Reading Scales in Biology, Scale B.

Purpose of test item:
To test student's ability to understand literature of a biological nature.

Directions: Read the paragraph carefully. Then read the first statement below the paragraph. If the idea it expresses is stated in the paragraph even though in different words, put a check mark in front of it. If the idea expressed in the statement can also be derived or inferred from the ideas in the paragraph place a check mark before the statement. Then read the other statements following and treat them likewise. Do not check statements which do not apply exactly.

18. The minimum essentials of a cell may be no more than a nucleus and cytoplasm, but rarely do we find cells so simple. A structure which must be present always at least functionally, whether structurally demonstrable or not, is a surface layer or membrane, a protective and discriminative film that bounds the cytoplasm. The cell membrane usually has a marked degree of toughness and elasticity and it serves to keep out of the cell substances that are inimical to life and to admit materials necessary for metabolism. It also shields the sensitive protoplasm from mechanical shocks and injuries. Even in apparently naked cells like amoeba, where there is no visible cell membrane, we know that there is a differentiated surface film that plays the role of a membrane; for naked protoplasm quickly yields to the cytolytic action of water. Is this membrane a part of the living cell or a mere dead product of the living cytoplasm? Emphatically we may say that the membrane is living, for it has all of the properties of a living thing. It is sensitive, conductile, contractile, and capable of growth and repair. It is, moreover, so highly individual in its make-up that it has the capacity of semi-permeability. A semi-permeable membrane is one that is permeable to solvents but more or less impermeable to certain substances in solution. This property is important in the life of the cell because protoplasm is a colloid solution and it is kept intact by the membrane, while water and the necessary dissolved food elements pass in and the dissolved waste products pass out. The cell membrane then is of prime importance in retaining the individuality of the cell and in presiding over the metabolic exchanges which form so large a part of the life of a cell.

1. Sensitivity, conductility, contractility, and capability of growth and repair constitute the properties of living cytoplasm.

2. All colloid solutions quickly yield to the cytolytic action of water.

3. The membrane of a cell is merely the dead product of the living protoplasm.

4. One of the functions of the cell membrane is to keep the protoplasm intact.

Applicability: While this test item makes no attempt to measure the desire to read scientific literature it does measure the ability to do so. The ability to comprehend what is being read is a tool of the desire to read. Without the tool the desire cannot get very far. Some measure of the ability to understand what is read should be incorporated into any credit given to achievement in the desire to read scientific literature.
SPECIMEN TEST QUESTIONS

from: Columbia Research Bureau Test in Physics.

Purpose of test item

Not stated in literature. The test in general is an informational one. Item shows a possible use in testing for the ability to apply principles.

Directions: Place a plus ( ) sign after any statement you think correct, and a minus ( ) sign after any that you think wrong.

Problem: A cube of glass 3 inches on each edge, with all faces polished, is placed over a black dot on a horizontal sheet of paper. The index of refraction of glass is \( \frac{4}{3} \).

91. the dot viewed from above and vertically, appears to be more than three inches below the upper surface of the glass. ..........( )

92. if light from around the dot is reflected at all from any of the surfaces, it is totally reflected. ..........( )

93. the beam from any point of the paper becomes a converging beam after passing through two parallel faces of the block. ..........( )
SPECIMEN TEST QUESTIONS
From: Malin Diagnostic Chemistry Test, Form A.

Purpose of items: to find weaknesses in the field of Chemistry.

First part is a table with one column containing many characteristics or properties listed. In the other columns the student has to check or reject these properties, for a list of several elements or compounds.

(A concise way of doing this and holding promise for adaptation.)

Part II Best answer multiple choice type.

Writing formulas
A. Write in the blank the correct formula of each compound, if correct, copy.

Sample NaPO₄. ........................... Na₂PO₄...

5. NaO ........................................
7. Al(NO₃) ..................................... etc.

B. Write the chemical name of the compound after each of the following:
Sample HClO₄. .......................... chloric acid.

10. HCl.............................
12. Na₃..........................
15. NaClO..........................

C. Write the correct chemical formula for each of the following:
Sample Sodium bromide ......... NaBr
19. Ferric hydroxide....................
20. Sodium perchlorate ..............

D. Balance the following equations:
22. AlCl₃ + NaOH ........................ Al(OH)₃ + NaCl

26. Na + H₂O ............................ NaOH + H₂

E. Complete and balance each of the following double decomposition equations:
27. FeSO₄ + NH₄OH ........................ + ................................
29. Zn + HCl .............................. + ................................

F. Solve the following problems, using the space provided for any calculations:
36. Find the water of crystallization in CuSO₄·5H₂O.
(Atomic weights: Cu—64; S—32; H—1; O—16)
Answer..........................

36. From the following equation, calculate how many liters of oxygen will be required to burn completely 20 liters of CO.

2CO + O₂ ................................. 2 CO₂

(All these are usual type questions.)
SPECIMEN TEST QUESTIONS.

From: Unit Tests, Caldwell and Curtis, "Science for Today"

Purpose of test item: To test the scientific attitudes held by the pupil.

Directions: The pupil was told to refer to pages 12 and 15 of Science for Today, by Caldwell and Curtis. There he would find a list of sixteen attitudes that characterize the truly scientific man. With these attitudes listed in front of him the student then had to decide which attitude fitted a certain situation. (This is really a matching question.) He then had to fill in the numbers of these printed attitudes as he saw fit.

5. "I know that our radio is the best kind made, because the man said so on the broadcast." ........................................

7. "The airplane was wrecked on Friday the thirteenth. I happened to think of the day and date just in time to change my plans so as to make my trip on the next day." ........................................

13. "Charles Goodyear spent many years in planning and making hundreds of experiments before he finally learned how to vulcanize rubber." ........................................

Applicability:

This test seems to test more a student's ability to judge the attitudes held by other persons or the best attitude to assume in a given situation, that is, the idealistic situation and solution. Doubt might be expressed whether these tests will test for the student's own attitudes in similar situations or not. The only test for this is to put the student in these situations either actually or vicariously and have him state his own attitude.

However, the test does bring to light the student's attention that certain attitudes are desirable, others not worthy of a fair-thinking person, and thus the form above may be used better as a teaching device.
From: Caldwell and Curtis, Test on Scientific Method (Tests for "Sc. for Today).

Purpose of test item: To test for the student's ability to analyze a problem into its component steps.

The paragraph which follows describes the activities of a scientist in solving an important and difficult problem. Each sentence is numbered. Write in the blanks following each of the phases of the scientific method outlined below, the number of the sentence or sentences which illustrates that phase or stage. Not every sentence that is numbered will illustrate a phase of the scientific method.

(1) A little more than a century ago, Daguerre, a French artist, was experimenting in order to discover how to develop a photograph negative.
(2) At length he was on the point of giving up in discouragement when one morning he was astonished to find that an exposed plate, which he had left in a cabinet the night before, had become developed. (3) He felt certain that the vapor of some chemical or combination of chemicals in the cabinet had developed the negative. (4) Could he succeed in discovering which chemical or combination had effected the change in the plate? (5) he decided to remove one bottle from the cabinet each day and to leave an exposed plate overnight with the remaining chemicals. (6) This would be a slow method, but one likely to solve the problem. (7) Days passed; the number of bottles slowly diminished, yet each morning he found the plate developed. (8) Finally one morning he found the plate undeveloped. (9) He had removed mercury from the cabinet the night before. (10) Mercury vapour must therefore be the chemical which had developed the negatives each time. (11) Daguerre therefore began to experiment with mercury vapour and freshly exposed plates. (12) The results were successful. (13) He had solved his problem. (14) He realized, however, that many further improvements in his process must be made before he could perfect a satisfactory process of photography.

PHASES OF THE SCIENTIFIC METHOD.

a. Locating and defining the problem.
b. Planning experiments.
c. Using controls.
d. Isolating the experimental factor.
e. Making careful observations.
f. Making inferences or drawing conclusions from the facts.
g. Making hypotheses from facts and observations.
h. Recognizing errors or defects in conditions or experiment
i. Evaluating conclusions in the light of the facts or observations upon which they are based.
j. Planning and making new observations, or checking experiments to find out whether certain conclusions are sound.
In this and the following two chapters an attempt will be made to adapt old forms, and devise new ones that will supplement those types mentioned in Chapter XIII in order to suit the needs of a testing programme based on the objectives of the General Science courses in high schools of British Columbia.

Each objective will be dealt with in order of importance and question forms will be suggested. The following is a general outline of procedure:

1. A re-statement of the objective in the order laid down at the end of Chapter XII.
2. Noting the percentage weighting given each by the averaged returns of the questionnaire.
3. An analysis of the objective for its specific aims or sub-objectives for the purpose of finding the major specifications to which a valid question in that field must conform.
4. The kinds of tests deemed suitable and valid for the specific objective.
5. Presentation of sample test items,
   (a) Purpose of test item stated when the objective is complex.
   (b) "Directions to students" for test item.
   (c) Test item.
   (d) Scoring; this is presented only for some of the newer types, and is not included where existing forms are modified only slightly.
Objective: To Acquire a Body of Knowledge in the Field of Science that Will Enable the Student to Interpret and Appreciate His Environment.

This objective of the General Science courses IV and V in the high school curriculum of British Columbia received the highest ranking and the greatest percentage weighting of all the objectives. The value when averaged gave this objective a weight of twenty per cent. This percentage weighting is probably lower than a true value might be, due to the conservative central tendency of averaging the rankings returned by the various teachers. Possibly a weighting of twenty-five or thirty per cent would be better. However, because it is the most important objective it should not be considered so important that it may exclude all others. It must be remembered also that the objective dealing with the development of the scientific method was ranked very close to this one.

There is a wealth of testing techniques suitable for the measuring of achievements in this field. They can be divided roughly into the essay or paragraph tests and the shorter objective tests. The various types have been segregated to some extent in the preceding chapter, and it could be seen from the analyses that some forms of questions fitted certain testing purposes better than others.

The first of these groups of questions is that of the essay or paragraph type and the many modifications. It is still much used, and probably will remain in use. Well prepared essay questions in this field still possess very valuable functions in the science courses. It is debatable whether the accepted mathematical type of question is any more removed from the charge of memoriter work than some of the essay-type questions unless special care is exercised. Students have memorized type problems in mathematics to pass tests just as
they have memorized material to be used in paragraph or essay questions. Changing the digits in mathematical questions in the attempt to form a new question is not a significant contribution to testing procedure. This "revised" question when answered by a student who has memorized the type of question is certainly worth no more than should be given for similar memoriter work in essay-type questions.

These questions have been modelled after the samples shown in Ruch and Rice's book "Specimen Objective Examinations", and are all suitable for testing procedures designed to measure the acquisition of knowledge. Thus they would be suitable to use in a programme under the first objective of the General Science courses IV and V.

Completion or Recall. (with varieties.)

The completion type of question is considered the most reliable according to Ruch and others who have worked on the problem of the reliability of objective questions. They vary in objectivity from the completely objective to the semi-objective.

A. Sentence Completion:

1. As we ascend through the atmosphere we find that the air pressure ...
2. The usual valence of calcium is ...
3. The gas used in nearly all forms of respiration is ...

B. Paragraph Completion:

When a copper wire forming a ... circuit is caused to pass through the field of force of a magnet a(n) ... is produced therein. This current lasts only as long as ... As soon as the wire moves in the reverse direction the current ... Coiling the same wire ten times to produce ten loops to cut the magnetic lines of force causes ... the current of one loop. Placing another magnet of equal strength adjacent to the first in cooperation with it and using the ten-looped coil will produce a current that is ...
than that produced by using one wire and one magnet.

C. Tabular forms of Completion Questions:

1. Complete this table of the properties of the various elements.

<table>
<thead>
<tr>
<th>Property</th>
<th>Mercury</th>
<th>Iron</th>
<th>Carbon</th>
<th>Nitrogen</th>
<th>Copper</th>
</tr>
</thead>
<tbody>
<tr>
<td>State under normal conditions of air, temperature; solid liquid or gas.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Gravity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action with oxygen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action with calcium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Describe briefly the following substances under the headings shown:

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Copper</th>
<th>Porcelain</th>
<th>Bakelite</th>
<th>Ebonite</th>
<th>Aluminium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action with electric current</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action with a static electric charge.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduction of heat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Gravity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3. Compare the following organisms on the bases given:

<table>
<thead>
<tr>
<th>Organism</th>
<th>Food</th>
<th>Methods of Locomotion</th>
<th>Obtains oxygen</th>
<th>Organs of sight</th>
<th>Body covering</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grasshopper</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clam</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frog</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grouse (bird)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolf (Mammal)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D. Another modification of completion questions can be derived from the mental testing diagrams of Cyril Burt. Instead of taking common things, such as the diagrams of a person's head, a ladder, and so forth, from which essential lines are omitted, diagrams of model pumps, actual pumps, plants or animals, chemical set-ups, and similar illustrations could be used for the student to complete the essential lines as in Burt's tests.

E. Chemistry makes much use of completion questions in the form of equation completion and formula completion questions. There are several variations of these.

1. What are the symbols for
   a. Calcium
   b. Iron
   etc.

2. What are the formulae for
   a. Hydrochloric acid
   b. Silver nitrate
   etc.
3. The symbol Cl represents the element Fe. 
4. The formula CaSO₄ represents the compound Na₂CO₃ represents the compound.
5. Complete the equation to show what compounds, elements, are produced or freed:
a. PbS + O₂ \rightarrow PbO 
b. AgNO₃ + NaCl \rightarrow +
c. FeS₂ + HCl \rightarrow +

True-false Question Types, With Variants.

The most commonly met variants of the true-false are the "yes-no", "right-wrong", and the "plus-minus" types of questions. These are so well known that anything more than a sample of each is unnecessary.

1. All stratified rocks are sedimentary. (True False)
2. Within elastic limits strain is proportional to stress. (Yes No)
3. Aluminium is a more common element than iron in the earth's crust which has been explored. (Right Wrong)
4. Photosynthesis occurs in all green parts of plants during only the sunlit hours.

A modification of these questions is to have the student provide the correct answer to all those questions which he marks wrong. These corrections are placed in blanks provided, as follows:

1. All stratified rocks are sedimentary. True False

Multiple Choice Questions.

A. Underline the best answer:

1. The substance which takes most heat to raise its temperature one degree Centigrade for each gram of substance is (aluminium, copper, water, alcohol)

2. Place in the brackets the answer that best completes the statement:
The mineral element that is obtained from the deeper strata of the earth's crust by the Frasch method is (phosphorus, carbon, copper, gold, sulphur, zinc).

This question form is not as desirable as the first or the following because it entails more effort from the student to do the mechanical work and it does not produce any greater mental effort, nor does it increase the validity or reliability of the test item.

3. Place in the brackets the letter (or number) of the answer that completes the statement in the best way:

   The process whereby each cell unites with food and oxygen to obtain heat is a. circulation, b. respiration, c. transpiration, d. inspiration, e. photosynthesis, f. digestion. (b)

4. Put a check mark (or a circle around) in front of the letter of the answer that best completes the statement:

   When any article is dropped from an altitude of about 15,000 feet through the atmosphere:

   a. the velocity of the body will continue to increase 32 feet per second per second until it reaches the ground.

   b. the body will fall at the same velocity during the complete fall.

   c. the velocity will be decreased as the object falls through the denser air near the ground.

   d. the acceleration will not be 32 feet per second per second because the atmosphere interferes.

5. Put in the brackets the number of the best statement to use in order to complete this sentence correctly:

   When common salt (NaCl) is used to treat hams or bacon either by rubbing it on them or by soaking them in a strong brine in order to preserve them the salt:

   a. preserves the food mainly by chemical change of the food.

   b. poisons the bacteria and the germs.

   c. causes an osmotic action that makes the bacterial cells lose too much water to permit them to live.

   d. when on the meat forms an anti-toxin against the germs and bacteria.
6. Mark "yes" or "no" to each statement below as you see fitting to do;
   (true, false)
   (right, wrong)
3. On weighing the reagents and products when silver nitrate and
   sodium chloride solutions that are hermetically sealed in a flask
   are mixed the results support the law.
   ( ) a. Energy can neither be destroyed nor created.
   ( ) b. When a mass of ions is injected into a chemical reaction using
       such ions the reaction goes in such a direction as tends to decrease
       them.
   ( ) c. Matter can neither be destroyed nor created.
   ( ) d. Gases in chemical reactions unite in whole integral ratios.

Multiple Response Questions.

In the questions of this type more than one correct response can be
included. They form a good antidote for the idea that students sometimes
hold that there is only one right answer to a problem. They can test a
series of relationships and demand a good deal of reasoning, as well as a
command of a broad field.

In the following questions more than one answer is correct. Underline
each answer that you think correct;
1. The following are good sources of protein (potatoes, eggs, bacon, milk,
   cheese, cakes, lean meats, carrots).
2. The production of iron in blast furnaces utilizes (sulphur, iron ore,
   tuyeres for hot air, reverberatory flame, coke or coal, and acid liming,
   limestone, a basic lining, phosphorus, tungsten, molybdenum).

Matching Questions

In making these questions it is imperative that they be homogeneous,
particularly so in General Science where a process of elimination of the
very diverse material that it is possible to put together will give the
student an answer which his actual knowledge does not warrant that he get.
Matching Questions (cont.)

In order to show the contribution of each man to science place the letter opposite his name in the blank in front of his particular contribution:

a. Scheele 1 . . . . . invented a reasonably cheap way to release aluminium from its compounds.
b. Hall 2 . . . . . . . invented the idea of molecules to explain chemical reactions and compounds.
c. Lavoisier 3 . . . . . invented or developed the microscope.
d. Dalton 4 . . . . . showed that gases increase in volume with temperature increase directly as their absolute temperature.
e. Charles 5 . . . . . discovered the nature of combustion and the importance of weighing all reagents and products in a chemical reaction.
f. Boyle
g. Galileo
h. Leeuwenhoek

In the above item #3 is not homogeneous, as all the other items are distinctly chemical, enabling a person with very superficial knowledge of the work to separate this item from the others. Also to use "Galileo" as one of the choices is next to useless as almost everyone recognizes his contribution as being in the field of physics. All choices must have discriminative value. When the list of blanks is short more extra choices must be provided than when the list of blanks is long in order to reduce the chance of arriving at answers by the process of elimination. When only five or six blanks are to be filled about ten to twelve choices should be given.

Rearrangement

In this type of question items must be rearranged in the proper sequence, either temporal when dealing with a process or development, or in some other manner. Some sequences have a definite starting point, while others are cyclic that can be started almost anywhere. For this latter group it is wise to give the first point of the cycle that you wish
to have rearranged in order to simplify the problems of marking. One question of each of these types is given.

1. Rearrange these items to give an outline of the contact process of making sulphuric acid.

1. drying the SO₂
2. absorption tower
3. Fuming H₂SO₄·SO₃
4. heating gases
5. roasting a sulphide to get SO₂
6. contact towers with catalyst trays
7. washing impurities out of gases


2. Rearrange these letters to show the complete course of the blood, beginning with the blood in the alveoli (small sacs) in the lungs of a human body.

a. left ventricle, b. cells, c. capillaries, d. left auricle, e. pulmonary arteries, f. right ventricle, g. lungs, h. right auricle, i. systemic arteries, j. air sacs or alveoli, k. pulmonary vein, l. systemic veins.

Starting with j the rearranged list would be: 1 . . ., 11 . . ., 111 . . ., 1V . . ., V . . ., VI . . ., VIl . . ., X . . ., XL . . .

Analogies

A fourth term that bears the same relationship to the third as the second does to the first is the usual question form of this type. The fourth term is to be supplied by the student. While apparently other terms like the second or the third may be left out for completion by the student all these questions can be rearranged to make that term the fourth. However a little variety is desirable.

1. platinum powder: converting SO₂ and O₂ into SO₃
   . . . . . . .: converting CO₂ and H₂O into C₆H₂O₆
2. longitudinal waves: sound
   transverse waves . . . . . . . . . .
Analogies (cont.)

3. feathers : birds

mammals.

Identifications

These questions seem to be a combination of matching and completion types. Ruch and Rice list them as distinct types.

1. In the exercise below you are to identify each substance as being an Element, a Compound, a Mixture, a Solution, or an Alloy by placing around each initial of these words a ring to classify each substance.

   quartz . . . . . . . . . . E C M S A
   beach sand . . . . . . . . . E C M S A
   air . . . . . . . . . . . . . . . E C M S A
   water . . . . . . . . . . . . . . . E C M S A

2. Mark the characteristics of Mammals with an M, of Vertebrates in general with a V, of Insects with an I, of Birds with a B, and use an A for a characteristic that applies to All, and N for one that applies to None.

   a. Diaphragm . . . . . . . M V I B A N
   b. Constant body temperature . . M V I B A N
   c. Nervous system . . . . . . . M V I B A N
   d. Jointed backbone . . . . . . M V I B A N
   e. Gizzard . . . . . . . . . . . M V I B A N
   f. Spiracles, trachea . . . . . . M V I B A N
   g. Milk glands . . . . . . . . . . . M V I B A N
   h. Fused or rigid bones, porous . M V I B A N
   i. Ovaries and spermaries . . . . . . M V I B A N

Reproduction from Memory

This is one of the last types listed by Ruch and Rice, and is very well known.

1. Write balanced equations to show what happens when sulphur is ignited in an atmosphere of oxygen; in an atmosphere of nitrogen.

2. Give Ohm's Law and explain the function of each factor in it.

3. By means of a balanced equation explain the process of photosynthesis in a simple form.
Deductions from Premises

This example comes directly from Buch and Rice's book "Specimen Objective Examinations", page 20. This type of question seems to reduce memorizing to a minimum and puts a premium on the ability to use data provided and on the understanding of the basic principles.

1. The following formulae are correct:

\[ K_2O, \ HCl, \ H_3PO_4, \ ZnO, \ H_2SO_4, \ CaO, \ KOH, \ KNO_3, \ Cr_2O_3 \]

From these facts calculate the valence of each of the following:

Chlorine ........ Phosphate radical ........
Potassium ........ Calcium ........
Hydrogen ........ Nitrate radical ........
Zinc ......... Chromium ........

Theoretically the questions above should demand active thinking in order to solve the problem. The examples chosen are all soon memorized by the average student. To avoid this difficulty stranger material for testing in this manner must be chosen, say the rarer members of these families, or use complete abstractions in \( x \) and \( y \).

Computations

These have been present in the educational systems from time immemorial. In these the emphasis is on the mathematical science. Because of their tendency toward the pure abstraction of scientific thought and reasoning it should not be assumed that these are the only kind of questions that test for scientific ability.

1. What weight of ammonium chloride will be required to produce 10 litres of ammonia measured at 23°C and 751 mm. pressure? (Chemistry Supplemental Examination for Matriculation, B. C. August, 1938).
2. What force will be required to balance a weight of 200 lbs. placed one foot from the fulcrum of a second class lever five feet long, ignoring the weight of the lever itself?

3. Yellow peas were crossed with green peas and all the resulting $F_1$ were yellow peas. If the $F_1$ flowers were all self-pollinated and the seed sown the next year to produce 12,000 seeds altogether, what kinds of seeds would they be and how many of each kind would be found?

To Develop the Ability in the Use of the Scientific Method: e.g.,

a. To make accurate observations and to record them systematically.
b. To draw valid conclusions.
c. To suspend judgment until sufficient evidence has been obtained.
d. To develop a critical yet tolerant attitude towards new ideas.

This objective has been ranked second with a percentage weighting of 18%. The objective naturally received a heavy poll, together with a considerable unanimity of opinion. It is one of the two major objectives, hence testing achievements toward this goal should be done with extreme care. The introduction of tests to measure the ability to use the scientific method would appear to be a decided innovation in almost all the schools of British Columbia.

The Science Revision Committee has sub-divided the objective into the four parts as above. Some authors suggest more steps or sections in the scientific method than these. Caldwell and Curtis in their textbook Science for Today seem to include some nine or ten, if statements scattered throughout the book are to be taken at their apparent value. Some of these seem to coincide. Then there are the more or less traditional five steps to follow in laying out a record of an experiment or the experiment itself: problem, apparatus, method, data or observations, conclusion. The objective of General Science IV and V omits very definitely
the first one which demands the ability to see that a problem is present and then to formulate it. This sub-objective should be included here and in the objective covering the development of actual experimenting.

No mention has been made of the isolation of the experimental factor. This is more or less a part of the general ability to recognize a problem. The use of controls or collateral experimental factors was not mentioned. These sub-objectives should be incorporated, whether the Science Revision Committee included them or not, for they are part and parcel of the scientific method.

Some authors and books include the ability to formulate hypotheses as a distinctly different function from that of drawing sound conclusions from data. The formulation of hypotheses and the drawing of conclusions seem to involve the same mental processes, the only difference being the degree of reliability of the data upon which each is based. No hypothesis is ever made but that it is based on some modicum of experience. The hypothesis is not a hypothesis by virtue of some different mental process but because the data from which it is drawn are extremely scant or untrustworthy. In situations like these scientists set up usually some "temporary working conclusion" to guide them in devising experiments to collect more data. This temporary conclusion is what is commonly called the hypothesis. Its purpose is to clarify the problem one step further, and in this does differ from the conclusion or law. It is nevertheless obtained by a similar mental process. The hypothesis is then set to work in other ways until sufficient data have been collected to warrant drawing conclusions that may be considered valid. Therefore, it is only the degree of difference in the reliability of the data that makes for a difference between
hypothesis and conclusion. On this basis little attempt to develop any
tests for hypothesis alone will be made, but attention will be limited
to the tests for the ability to formulate conclusions from data, and to
realize when data are too deficient to trust completely.

During the period when the main work of examining test items was
in progress the lack of testing techniques for the purpose of measuring
ability to use the scientific method became apparent. The only question
type that occurred occasionally was that of "deduction from data", and
which was practically always mathematical data. Such questions are known
commonly as the "math problems" in physics and are found to a lesser extent
in chemistry. With this big gap in testing techniques in view, work was
done devising means to measure achievements in this particular field. The
following question types were developed after much hard work.

The first items appear on the next page.

**Purpose of Test Items:**

To measure the ability of the student in making accurate observ­
vations and measurements.

To see if pupil can detect common parallax errors, can use the beam
balances accurately (that is make correct readings), and realizes
where the chief errors are most likely to occur in measuring liquids.
1. This problem arose in a meat store.

A lady customer, 5 ft. 7 in. tall, ordered two pounds of steak from a butcher whose height was about 5 ft. 1 in. The man put some meat on the scales, and after they came to rest said, "Two pounds." The woman disagreed, claiming that they were two ounces short in weight. The butcher leaned closer to the scales and declared that the meat actually weighed one ounce more than he first said. Position C is the customer's eye. Position P₁ the butcher's first. Position P₂ the butcher's second.

Answer These Questions on the Above Problem

1. Who was right? (butcher, customer, neither, both)

2. Why did the butcher on leaning closer maintain that the weight was actually greater? ...........

3. Use the letters to designate the reasons in the list below that you would use to support your decision in question 1. ...........
   a. The butcher because he knew that the scales were "fixed", or incorrect.
   b. The butcher, because he knew that the scales were quite accurate.
   c. Neither, because their eyesight might have been poor.
   d. Both, because it's all in the way you look at things.
   e. The customer, because she knew or suspected that the scales were "fixed" or incorrect.
   f. The customer, because she suspected that the butcher had placed his thumb on the pan to increase the apparent weight.
   g. Neither, because the true weight lies halfway between the two.
   h. The customer, because of the parallax of the butcher's eyes.
   i. The butcher, because of the parallax of the customer's eyes.
   j. If none of these, add your own reason ..........

11. In weighing some powder the person set the beam balances true first, then poured forty cubic centimeters of it on a thin piece of paper. This set was then placed on the left pan, and a 50 gram weight, a 20 gram weight and another 20 gram weight were added to the right. The
rider or marker on the beam at the top was moved toward the right until the scales balanced at the position shown. The rider had been placed previously at .8 in order to compensate for the piece of paper which was placed on by itself.

What is the weight of the powder measured out? . . . . . . . . grams.

![Figure XV](image)

111. You wish to measure 100 c.c. of water into a graduated vessel of 200 c.c. capacity. What are the three most important things to do in order to get exactly 100 c.c. measured into the graduate?

1

2

3

111. In reading the temperature registered on a thermometer being used to measure the temperature of melting ice, in a large glass beaker, care should be taken to follow the suggestions in the statements lettered:

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

a. Do not spill any water from the beaker as it will alter the temperature.

b. Place only the bulb of the mercury in the ice pack.

c. When reading the thermometer take it out quickly and very carefully and hold it near to the eye to be sure.

d. See that the bulb and part of shaft with mercury in it are buried in the ice pack.

e. Place the bulb of the thermometer one-half inch above the ice surface.

f. In reading the temperature have the eye and mercury top on the same level.

g. Put ice and thermometer set in the refrigerator.

h. Read the temperature on thermometer while the thermometer is in the ice pack.
Scoring

Question #1, one point per blank.
Question #2, three points, one point for the correct gross total (involving the rider value and weights; one point for deducting the weight of paper; and one point for correct net total.
Question #3, one point each for: horizontal base, on eye level, lower surface of meniscus curve.
Question #4, one point each for d, f, h.

Questions 11 and 111 were tested by the investigator to see if the paper test would actually test laboratory techniques. On the conclusion of the experiment (Exp. IV in Appendix) it was seen that the two methods of testing are not testing the same things, as the correlation is low. Moreover, the paper test seems the harder (except for measuring volumes of liquids), for in this and other experiments the score on the actual manipulation was higher than on the paper score. (It is quite conceivable, however, that this condition might be reversed easily if the pupils did not obtain much training in laboratory work. The students of the classes tested had done a good amount of laboratory work in the previous years).

Purpose: To test the student's ability to form hypotheses (a form of deduction from data).

1. In diamond drilling along the base of the Rocky Mountains in Alberta the geologists found the cores yielded series of rocks in two groups as shown in the diagram. What hypothesis would you advance to explain this? Explain what you would do to confirm or destroy this hypothesis.

![Test Port Hole](image)

Sedimentary Strata
Metamorphic Zone
Sedimentary Strata
Metamorphic Igneous (Plutonic)

Figure XVI
II. You have been called in to find some solution to a serious traffic problem occurring at a five-street junction, without any special traffic rules for that corner. What would be five hypothesis that you might advance in solving the problem, and that you would expect to test out experimentally? (Too many accidents are occurring at this junction.)

a. ........................................

b. ........................................

c. ........................................

d. ........................................

e. ........................................

Scoring: For each of the above problems there are possibly several sensible hypotheses. For the first one these hypotheses may be advanced;

1. overlapping hypothesis advanced, 2. the repetition of similar histories for that area in two different periods of time, 2. cataclysmic explosions superimposing one series on top of itself, etc. Each sensible hypothesis would have to be counted. A series of questions should be given to attempt to reduce subjective evaluations to a minimum, that is, sampling should be large.

For the second question many hypotheses might be advanced such as;

1. too great a speed as causing the accidents, 2. poorly kept vehicles, 3. too much extremely slow traffic in a heavy stream, 4. types of traffic too mixed, pitch of hills, 5. road surfaces, 6. business of neighborhood.

Each hypothesis within reason would have to be accepted, and a complete list of responses might not be built up for a considerable time, because of repeated additions. One point each could be given.

Purpose of Test Items:

1. To test pupil’s ability to organize a problem into logical steps.

2. To test the pupil’s ability to recognize the parts of a scientific problem.

Directions:

The group of sentences which follow have been taken from an account of a rather famous scientific problem that was solved. The statements have been put in a different order, and then lettered. Read them through carefully, trying to piece together the logical sequence of events. At the end of the series are questions to be answered.
Many triangulations and calculations were made by both investigators.
On the night of September 23, 1846, Galle turned his telescope to the point in the sky where the new planet was predicted to be, and found it.
Something under certain conditions produces other actions.
For the first forty years after its discovery Uranus followed the correct path like all other well-behaved planets.
Both Adams and Leverrier attacked the problem without the other's knowing it.
The heavenly body acting on Uranus must be farther out in space.
The new planet was named Neptune.
The planet Uranus in 1820 was slightly out of place and expected time, being ahead and out further.
The theories upon which the position of Uranus for 1820 was predicted were entirely erroneous.
The planet Uranus is being pulled out of its normal path by another heavenly body farther out, and this heavenly body is unknown up till now.
The new heavenly body must lie where the lines of triangulations and the calculations coincide.
One of the greatest triumphs of mathematical genius occurred in 1846.
The laws of mechanics and gravitation were proven wrong for the first time by a concrete example of different actions possible in heavenly bodies.
The heavenly body acting on Uranus must be nearer the Sun than Uranus.
By 1840 Uranus was so far out of position that the discrepancies were intolerable to a meticulous astronomer who put faith in the calculations.
Why was Uranus not in the positions ascribed to it by the calculations?
Uranus is out so far in space that its surface temperature is probably below zero by 200°F. during the middle of the day.
Many involved mathematical calculations based on the laws of mechanics and gravitation showed that the new heavenly body must be inside of Uranus' orbit.
Ten years later the differences in time and position were still greater, so that Uranus was an appreciable distance outside its orbit.
It is known that any two heavenly bodies tend to draw each other together by virtue of their gravitation.
Some other heavenly body was acting on Uranus to pull it out of its calculated position.
The index of reflection of Uranus is very low so that the observations were not reliable.

By using the letters in front of each statement re-arrange them to include only the necessary statements and to form a series that gives the correct sequence of events and actions.
Be sure to use the letters of only the statements that serve a direct function in the problem.
The scoring of questions of a cyclic nature is always rather difficult. What are you to do when a student slips up on one step early in the cycle but all the others are in order but out of their correct position? I would suggest that for any runs or sequence that the score for that portion be \((n-1)\) where "n" stands for the number of items in correct sequence for that portion. One slip will automatically drop the score two points.

This exercise could have the questions asked in another way. Ask specific questions about certain statements as:

1. What is the use of statement "j" in working with the scientific method?
2. What is the function of statement "s"? (problem, hypothesis, data, conclusion, unnecessary).
3. What is the function of statement "k" in solving the problem?

This form of question makes for greater ease in marking for no worries of evaluating sequences are present. On the other hand it loses considerably its validity as a test for the ability to organize the steps of a problem.

Another form of questions based on this exercise would be:

What statements of those lettered above form the:

1. Problem
2. Examination of older theories
3. Experimentation
4. 

This type of test would serve nicely in testing for knowledge of the various steps in the scientific method, but it loses somewhat validity as test of organizing ability. To organize, the student must
really do the organizing of the whole question. Questions that present him with a partial organization of the preceding two types invalidate the purpose slightly. Question might have the first few steps already organized by letter, or the middle two or three, or the last. This would seem to invalidate the test less as a test of organizing than the second and third types do.

Purpose of Test: To test ability to interpret data.
To test ability to draw valid conclusions.
To apply the generalization accurately to applicable new situations.

1. The following represents the results of some research into the effect of velocity on the air resistance of an automobile travelling at various speeds.

DIRECTIONS: Placed after the data and graph are some statements and conclusions. You are expected to read all data carefully and to make all decisions on the basis of what is presented to you here. Examine each of the statements below, decide which of these judgments fits it best, then insert that letter in front of the judgment in the parentheses in front of the judgment.

a. A SOUND CONCLUSION based on evidence only.
b. A conclusion CONTRADICTORY to evidence shown.
c. QUITE PROBABLE, but evidence or data does not go that far to show.
d. NOT LIKELY CORRECT, but evidence does not go that far to prove definitely.
e. NOT RELATED to the experiment, INAPPLICABLE.

![Figure XVII](image-url)
COMMON VELOCITIES:

1. Automobiles 25-45 miles per hour.
2. Speed limits in B. C., 30 miles per hour.
3. Launches, 7-15 miles per hour.
4. Coastal steamships, 15-25 miles per hour.
5. Aeroplanes, 90-200 miles per hour.

Hull—that part of a boat floating in water, and supporting the superstructure.

Stream-lining—Rounding the corners, and tapering the shapes, of bodies in motion to reduce the vacuum behind, or the "after-drag".

( ) 1. Streamlining the body of automobiles would be of little importance for machines that travel within the legal speed limits.
( ) 2. The resistance of air increases proportionately to the velocity of the car.
( ) 3. Streamlining the hull of a launch would be of very little assistance in saving fuel, or in gaining speed.
( ) 4. Driving automobiles at higher speeds produces increasingly greater air resistances.
( ) 5. The velocities in air above which resistances increase very rapidly are above 40 miles per hour.
( ) 6. The measurements are not accurately made because the graph is not curved evenly.
( ) 7. The streamlining of aeroplanes is really unnecessary, it being mainly a type of artistic designing.
( ) 8. Streamlining the hull of steamships would not reduce resistance very much.
( ) 9. Streamlining the superstructure, upperdecks, etc., of a steamship will not reduce the air resistance to any great extent, and therefore is not necessary.
( ) 10. The weight of the air affects the resistance offered the car.
( ) 11. Mileage per gallon of gasoline used by an automobile would be less at 60 miles per hour than at 30 miles per hour.
( ) 12. Mileage per gallon of gasoline used by an automobile would be greater at 50 miles per hour than at 25 miles per hour.

SCORING: One point each, and the total of correct responses.

Purpose of Test: To test the ability of student to arrange data in a systematic manner, more particularly this time by use of graphs.

2. To test ability to interpret data.

3. To test the ability to suspend judgment in the face of insufficient facts.

4. To predict results on the basis of data.
DIRECTIONS: On the graph paper record the data given below, then use your graph to answer the questions asked.

DATA

In finding how well a certain salt dissolved in water at various temperatures a laboratory worker observed these facts:

Salt A:
At 20°C. .... 31.81 grams, dissolved in 100 grams of water
45°C. .... 73.4 " " " " " " " "
0°C. .... 13.3 " " " " " " " "
5°C. .... 15. " " " " " " " "
40°C. .... 63. " " " " " " " "
60°C. .... 110.75 " " " " " " " "
10°C. .... 17.2 " " " " " " " "
30°C. .... 46. " " " " " " " "

Salt B:
At 0°C. .... 3.8 grams dissolved in water per 100 grams of water.
At 5°C. .... 4.2 " " " " " " " "
90°C. .... 47.6 " " " " " " " "

1. How much of salt A will be dissolved in 100 grams of water at 70°C?

2. How much salt per 100 grams of water will be dissolved at 15°C?

3. How much salt of sample A will be dissolved in 500 grams of water at 50°C?

4. How much of salt B would dissolve in 100 grams of water at 55°C?

5. How much of salt B would dissolve in 100 grams of water at 100°C?

6. How much of salt B would dissolve in 400 grams of water at 75°C?

7. At 50°C. which salt dissolves better in 100 grams of water?

8. What relationship exists for salt A between quantity of salt dissolved and temperature?

SCORING: The first three questions can be scored perfectly objectively above a slight tolerance that should be permitted for instrument errors, pencil lines, etc. (as this is not a method whereby the student can obtain
the exact mathematical results). A tolerance of, say .2 grams, would be permissible. Each correct answer would be given the one point. For questions 4, 5, 6, 7 the best answer, and really the only one correct, is that the "data are insufficient", or "I don't know". Any results based on a straight line graph must be thrown out for just enough evidence is presented to show the careful worker that this is not of that type. However, some of the more ambitious, taking a hint from the graph of A will attempt a curved graph. In cases like these the examiner will be forced to go back to his original data to find if the responses should be accepted. Really they should not, as this is a test to see if the student realizes that he has not sufficient data. For question 8, answers should be of this type and value; increases more than proportionately to temperature, 3; increases, but not evenly, 2; increases with temperature, 1; no answer, 0. A better evaluation for questions 4, 5, 6, 7, would be to give 3 points for "insufficient data" or equivalent; 2 points for responses within tolerance limits set; one point for no answer; and zero for all others.

The actual graphs should be marked, too, say two points for a graph that is correct ("the arranging of data"), one point for a full graph that is rather careless, others zero, for salt A; for salt B the graph should run between the first two points with a slight curve, to get two points; a straight line from the second point to the third could well be given a minus mark of one.

Purpose of Test Item: To test the ability of student in drawing valid conclusions; in suspending judgment in the face of insufficient facts; and in identifying the variable factors in a complex.
**DIRECTIONS**: Read carefully the data presented, and the statements below that. Each statement you are to judge on the basis of the five standards or judgements listed below and lettered. In the parentheses in front of each statement record the letter of the judgment that best fits that statement.

The Late Blight of Potatoes is caused by a fungus parasite (*Phytophthora infestans*) that grows throughout the potato plant, particularly the leaves, and producing death in those parts infected. When infection is severe the whole plant may be killed and killed very rapidly, seemingly overnight. This is what happened in Ireland in 1845 to cause the Irish Famine, and it resulted indirectly in the death of thousands of Irish and the emigration, to America mainly, of about a million people. In order to try to control such a disease many experiments have been done. Below is the table showing the data from one of these experiments studying the factors that affect the disease.

Plots of land 1/10 acre each were marked out on very good soil that had been under proper cultivation for many years. Each year new plots were used and a rotation developed. Three plots were chosen each year; in one the potatoes were sprayed on a very rigorous schedule with Bordeaux mixture which forms a copper hydroxide film over the leaves; another plot of potato plants was sprayed on an equally vigorous schedule by using lime sulphur spray of the proper concentration (as found by other experiments), while in the third plot the potato plants were not sprayed at all. The blight attacks very severely during July and early August if conditions are right for it.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>WEATHER DURING JULY, rain, sky</th>
<th>AVERAGE TEMP JULY</th>
<th>YIELD BORDEAUX</th>
<th>YIELD LIME SULPHUR</th>
<th>YIELD UNSPRAYED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>moist, alternating rain, warm</td>
<td>56° F.</td>
<td>3,255 lbs.</td>
<td>2,982 lbs.</td>
<td>2,070 lbs.</td>
</tr>
<tr>
<td>2.</td>
<td>dry in June, occasional showers July</td>
<td>62° F.</td>
<td>3,409</td>
<td>3,024</td>
<td>1,207</td>
</tr>
<tr>
<td>3.</td>
<td>generally fine &amp; clear, little rain</td>
<td>73° F.</td>
<td>2,810</td>
<td>2,643</td>
<td>2,725</td>
</tr>
<tr>
<td>4.</td>
<td>showery generally</td>
<td>59° F.</td>
<td>3,229</td>
<td>2,956</td>
<td>2,046</td>
</tr>
<tr>
<td>5.</td>
<td>moist, alternating rain &amp; warm</td>
<td>64° F.</td>
<td>3,156</td>
<td>2,740</td>
<td>1,477</td>
</tr>
<tr>
<td>6.</td>
<td>fine, few showers</td>
<td>71° F.</td>
<td>3,566</td>
<td>3,179</td>
<td>1,697</td>
</tr>
<tr>
<td>7.</td>
<td>fine, clear no rain</td>
<td>74° F.</td>
<td>2,892</td>
<td>2,459</td>
<td>2,706</td>
</tr>
<tr>
<td>8.</td>
<td>fine, clear, no rain</td>
<td>70° F.</td>
<td>3,276</td>
<td>2,933</td>
<td>3,162</td>
</tr>
<tr>
<td>9.</td>
<td>very rainy</td>
<td>62° F.</td>
<td>3,582</td>
<td>3,092</td>
<td>1,679</td>
</tr>
<tr>
<td>10.</td>
<td>cloudy not much rain</td>
<td>66° F.</td>
<td>3,217</td>
<td>2,769</td>
<td>1,946</td>
</tr>
</tbody>
</table>
In front of each of the numbered conclusion place the letter of the one statement above which you think describes it best.

- a. SOUND CONCLUSION based on evidence only.
- b. a conclusion CONTRADICTORY to evidence shown.
- c. QUITE PROBABLE but evidence or data do not go that far.
- d. NOT LIKELY CORRECT but evidence does not go that far.
- e. NOT RELATED to experiment or INAPPLICABLE.

Conclusions:

1. Any fungus-killing spray gives equal and adequate protection. (a)
2. Hot, dry weather hinders the development of Late Blight disease. (b)
3. Rain and moisture play no part in the action of the parasite in causing the disease. (c)
4. All data of yields are invalid because the areas of comparison are not equal. (e)
5. Bordeaux mixture seems actually to increase the yields. (a)
6. The best soil was chosen for the Bordeaux sprayed plots. (a)
7. Sufficient data have been obtained to draw valid conclusions as to worth of spraying. (a)
8. Lime-sulphur spray reduces crop yield of potatoes in some way. (a)
9. Late Blight is worst in the hottest years. (a)
10. The colder the temperature the better the parasite grows, killing more potatoes. (a)
11. Lime-sulphur spray of the concentrations used protect the potato plants from the parasite. (a)
12. Because the spraying costs on the average 60¢ per plot and potatoes average only 59¢ per 100 pounds it did not pay financially to spray. (a)

Purpose of Test Item: To test the ability to detect principles at work as shown in the data (drawing valid conclusions) and to apply these principles to new suitable situations.

(Question type suggested by Professor C. B. Wood)

Directions

This table presents data obtained by measuring current and the strength of the magnetic forces produced when electricity was sent through helices or coils of differing numbers of turns. Three trials were uncompleted; you are expected to fill in these blanks with your calculations based on the evidence.
Scoring: The scoring is very objective for this, simply one point for each correct response only.

The ability to make accurate observations and to organize these data and record them neatly are achievements that can be evaluated directly from the student's note books, or laboratory books to obtain a fairly satisfactory measure. Perhaps this may be a good interim procedure until tests that are more valid than the marking of books can be devised. The marking is more valid when restricted to organizing.

In marking the laboratory books the best plan is to examine all books and separate them into seven groups accordingly from best to poorest. It is advisable to check the placement of books in any category, for a few of the first ones examined possibly may have been misplaced when viewed in light of the fuller experience. These can then be allotted letter grades, or points, according to the method in use of recording achievements.

<table>
<thead>
<tr>
<th>Turns of Wire in Coil</th>
<th>Current in Amperes</th>
<th>Magnetic Strength in Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>100</td>
<td>$\frac{1}{5}$</td>
<td>50</td>
</tr>
<tr>
<td>200</td>
<td>3</td>
<td>600</td>
</tr>
<tr>
<td>700</td>
<td>$\frac{1}{2}$</td>
<td>350</td>
</tr>
<tr>
<td>800</td>
<td>$\frac{1}{10}$</td>
<td>80</td>
</tr>
<tr>
<td>600</td>
<td>$\frac{1}{3}$</td>
<td>a...</td>
</tr>
<tr>
<td>800</td>
<td>b...</td>
<td>200</td>
</tr>
<tr>
<td>c...</td>
<td>$\frac{1}{4}$</td>
<td>100</td>
</tr>
</tbody>
</table>
CHAPTER VI

FORMS OF QUESTIONS FOR THE OBJECTIVES OF "INTERMEDIATE" VALUE.

In this chapter are grouped the test items suitable for testing achievements of the objectives of intermediate value. These objectives obtained a weighing of eleven to ten percent, which marks them off abruptly from the two preceding but only slightly from those which follow.

To Develop Resourcefulness and Adaptability to New Conditions:

This objective was ranked third with a weighting of eleven percent.

Judging by the returns on some of the questionnaires several persons think that this objective is debatable on psychological grounds. Four expressed the opinion that it is psychologically impossible to develop resourcefulness because it is a characteristic of intelligence, or on the basis of Spearman’s theory one of the specific factors.

These comments set the investigator to examining the theories of intelligence again just to check up. It is not intended to go into a lengthy discussion of the theories of intelligence at this stage and to use ideas deducible from the theories but rather to deal with direct experimental evidence on the problem of testing and developing resourcefulness.

It might be said in defense of this objective that one of the outstanding characteristics of a resourceful person is his adaptability. While doubtless adaptability is in a measure based on intelligence, or more correctly upon the degree of responsiveness of an individual, it is

not synonymous with intelligence. It is rather on attitude of mind which can be developed in a person. A very intelligent person can be unadaptable even in these fields. Adaptability is developed by experience and correct attitudes. The degree of development is of course limited by the degree of native responsiveness of the individual. Resourcefulness may be characteristic of persons. The great American nation has been derived primarily from British stock. It is doubtless a very resourceful nation, possibly much more so than are the British. This difference may be attributed to their need to adapt themselves to new and different surroundings. This need is not evident in the lives of the British for whom conditions are still similar to those of past days and are not changing rapidly.

It seems quite possible to make out a strong case to show that resourcefulness can be increased in any person by creating the proper attitude of an open mind, by showing that problems often have more than one solution or one way of solving them, by improving powers of observation of details around one, and by teaching students to arrange their observations in some orderly or systematic manner. All these processes can be developed by good training in science. Thus it seems in order that this objective be retained. There is evidence, too, to show that resourcefulness may be developed.

One experiment carried out by Beauchamp and Webb tried to test the resourcefulness of high school students. These investigators reported that there was very little correlation between resourcefulness as they measured it and intelligence as measured by two well known general intelligence tests. The Otis intelligence test gave a correlation of $0.21 \pm 0.09$

1. For examples of these see p. 133
with resourcefulness, and the McCall Multi-mental test .04 .09. The
Yerkes-Ross Adolescent-Adult scale correlated with the resourcefulness
tests to the extent of .42 ± .08. Therefore, they conclude, the intel-
ligence tests are not measuring the same things as their resourcefulness
tests were, at least to any great extent. It might be pointed out here
that resourcefulness correlated with the achievement in physics, which was
measured at that time in the rather bookish manner, to the extent of
.14 ± .09. These investigators go on to say that if resourcefulness is
really a desirable factor of a pupil's ability then we are not justified
in measuring him only by tests that do not measure this quality. Growth
in resourcefulness was reported also.

Thus it seems that the inclusion of this objective in the Science
courses IV and V is quite justified. Moreover it has been judged quite
important by many teachers with a fairly good degree of unanimity.

Some examining of the term resourcefulness is necessary in order to
see just what is implied. The average situation demanding resourceful-
ness is one that has a new element in the problem. We say that the person
who is able to solve such problems with new elements is resourceful, the
more so if he can solve them quickly. In solving them he usually makes
use of the very simple things that are around him. It is necessary not
to confuse this characteristic with ingenuity; the two may be closely
related, but ingenuity has more inventiveness to it. In making his
solution to the new complex the person also makes much use of experience.
If the foregoing is correct then these seem to be the characteristics of
resourcefulness: to realize the difficulty or problem, to survey the
materials that can be used in solving it, the manipulation of these ar-

Mathematics May 1927, pp. 417-418. (Vol. xxvii)
ticles, the call upon experience or memory for parallels, and the calling in of intelligence. What can be done to develop resourcefulness then? Students can be trained to identify problems; they can be taught systematic surveying of materials that might be useful; the provision of a wealth of experience, direct or vicarious, will prove of inestimable value; they can be trained to manipulate apparatus. So from this it seems that much can be done to develop resourcefulness, and thus the testing programme should follow the teaching. Perhaps it might be better to say that the knowledge that tests can be devised for this purpose should encourage the teacher to include training for this in his programme.

Any test of resourcefulness must place the person in a new complex of circumstances. There are many problems in the daily teaching experience that call for resourcefulness. To enumerate many here would be to rob such incidents of the element of the new. Many laboratory activities and techniques that form part of the regular work in higher years have real elements of resourcefulness in them when applied to the lower grades. Care must be taken when choosing from these that they are not entirely beyond the experience of the student. The type of test administered by Beauchamp and Webb demanded actual things to work with, although Beauchamp and Webb did divide their group into two, and alternate paper response and actual manipulation between the two sections.

They provided a booth for each student to use. In the written response type all the equipment and the problem, which was printed on a card were uncovered on a given signal; whereupon the student had to see what the problem was, survey his apparatus, then write briefly his procedure. For the actual manipulation group also the problem was printed on the card, but the actual solution demanded the use of the apparatus. A time
limit of seven minutes per item was deemed sufficient. Each student was asked also to state whether or not he had done something similar to the test item before. Out of the 1,309 successful accomplishments 896 "were based on some modicum of experience. A correlation of .60 between experience and resourcefulness shows the benefit of experience in any resourceful act.......Resourcefulness is not mere repetition of experience but it is founded on it." The scoring was entirely in the hands of the supervisor who watched the actual manipulations. Successful responses only were counted and no fractional credits given.

A few of the items from Beauchamp and Webb's list will be included here:

**FORM A**

1. Given: Bunsen burner fastened down; gas supply; matches; short rubber tubes (too short to reach from jet to burner); glass tubes (of slightly smaller external diameter. Required; To light the Bunsen burner without moving it.

3. Given: Two bottles of odd shape, nearly the same size; pan of water. Required; To find which bottle holds the more.

10. Given; A string one yard long; scissors. Required; To secure accurately a strong six inches long.

11. Given; Flask; one-holed stopper; funnel; water. Required; To prove that air occupies space.

**FORM B**

1. Given; Basket ball bladder; balances; weights; string. Required; To prove that air has weight.

Without going further one is tempted to say that the last two activities in particular are included in science tests and science courses in many places. This probably does not matter in the least, providing that they have not been done already in the class or grade which is being tested.
Some types of questions that seem to be answerable only on paper and having the characteristics that demand resourcefulness are:

1. Given: Two steep hills and a moderately narrow valley; a watch with second hand; rifle and shells; a person stationed or resting about half-way up one of the hills.
   Required: To find the distance across the valley at the altitude of the person.

2. Given: A fish that has just been caught; small trees and willows near by; "flies"; hooks; spinners; basket; string or line; knife; lead-sinkers 1 oz to 4 oz.
   Required: To find the weight of the fish.

3. Given: A small fire from a short circuit in your car; out on a country side road far from help; no fire extinguisher.
   Required: To put out the fire immediately.

Others that seem to demand actual materials to be present in order that the student can make the necessary thought connections are:

1. Given: some empty halves of peanut shells; sand or soil; salt or sugar.
   Required: To demonstrate the process of osmosis.

2. Given: An angel-cake tin; and three other pans of slightly larger diameter, alkaline water (or some other with chemical impurities) fire or Bunsen burner.
   Required: To obtain water chemically pure in order to put it in storage batteries.

3. Given: Potatoes; pan of water; knife; sugar.
   Required: To prove that water passage from cell to cell occurs by osmosis.

4. Given: A tack at the bottom of a half-inch hole six inches deep and drilled into a piece of wood that cannot be turned over; nails and spikes up to seven inches in length; a dry cell; and thin silk-covered copper wire.
   Required: To get the tack out so that drilling may be resumed, and the wood not injured.

5. Given: A shovel with a cracked handle but not separated; a very heavy rock (about 250 lbs.) in a hole two and one-half feet deep; dug in a garden.
   Required: To get the rock out of the hole, assuming that it is much too heavy to lift.
To Acquire Knowledge Which Will Contribute to Public and Personal Health

This objective was ranked fourth and received a weighting of ten percent.

There seems to be no confusion as to the intent of the objective. A science teacher is directed by this objective to see that the required knowledge or information is given when the science course approaches health topics, or when principles are involved which are transferable to health and safety studies. While there exists a special series of Health courses in British Columbia schools no one expects the science teacher to teach the entire health course. There will be many places where the Science and Health courses meet or overlap. These linkages should be made very evident to the student.

The objective emphasizes the acquisition of knowledge that should be put into active use by the student to safeguard his own health and that of the community. To see that this information and training is given and to try to see that it becomes effective is the duty of the teacher. The testing for the information acquired will be relatively easy. To test to see how socialized health attitudes have been developed will not be too difficult. The common forms of questions can be used for these. To measure the transfer of information into active life of the community would be a big task beyond the power of one teacher to do.

MULTIPLE CHOICE: Underline the best answer.

1. During the winter the air taken into the hot air heating system of any building and warmed only becomes: a. too humid, b. too warm, c. slightly higher in humidity, d. slightly lower in humidity, e. too dry.

2. Air that is still and very humid hinders considerably the process of a. transpiration, b. respiration, c. perspiration, d. circulation, e. digestion.
3. The best method of illumination to use where much reading will be done is: a. direct, b. totally indirect, c. semi-indirect, d. totally direct.

MULTIPLE CHOICE QUESTIONS: Place the letter in front of the term in the blank numbered in front of each statement so that a closely related pair of ideas results.

a. filtering  
1. added in minute quantities to kill germs that may be in the drinking water of many large city systems.

b. drinking water  
2. greatest dangers from water supply occur.

c. septic tanks  
3. single drop in a tumbler or glass full of water usually kills all germs.

d. sewage  
4. does not remove germs from water.

e. boiled water  
5. disease spread chiefly through water and milk supplies.

f. typhoid  
6. sufficiently pure if it contains nothing unwholesome.

g. chlorine  
7. typhus fever occurs.

h. summer vacations  
8. typhoid fever occurs.

i. fluorine  
9. typhus fever occurs.

j. diphtheria  
10. typhus fever occurs.

k. tincture of iodine  
11. typhus fever occurs.

l. boiling  
12. typhus fever occurs.

m. typhus  
13. typhus fever occurs.

n. spring  
14. typhus fever occurs.

COMPLETION OR RECALL: fill in the blanks in the statements in order to make a complete accurate statement.

1. The souring of milk is caused by .........................

2. In short sightedness the image falls ..................... the retina of the eye.

3. The invention based on the use of electricity and vacuum tubes that is of most benefit in surgery and diagnosis is .........................

ANALOGIES: complete the analogies below with a word or statement that makes the same or very similar relationship between the second pair of words as exists between the first pair.


3. flies : typhoid : : fleas : ............... 

4. P₅O₅ (in saliva); starch : : pepsin : ............... 

5. wood alcohol; grain alcohol : : ethyl : .....................
IDENTIFICATIONS; classify these afflictions as, a. communicable diseases 
b. non-communicable diseases, c. an injury from external causes, d. a hereditary defect. When you have decided to which of the above group each affliction belongs write the letter of that group in the appropriate blank.

1. concussion of the brain
2. six fingers on each hand
3. tuberculosis in cattle
4. potato scab
5. diabetes of man
6. winter killing of bark of trees
7. hog cholera
8. scurvy
9. typhoid
10. wheat rust

DEDUCTIONS FROM PREMISES, (drawing valid conclusions).
Read the paragraph as carefully as possible then answer the questions at the end of the paragraph. Be careful to make no statement that is not deducible from the quotation.

"Metchnikoff, a great Russian biologist, observed that the Bulgarians are an unusually long-lived people. In trying to discover the reason for this he searched for something in their living conditions which was different from the living conditions of all other peoples. He found, among other things, that the Bulgarians drink more buttermilk than other peoples. He therefore concluded that the drinking of buttermilk was responsible for the long lives of the Bulgarians, and that if other peoples were to adopt the custom of drinking as much buttermilk as the Bulgarians, they would live as long as the Bulgarians.

1. What causes long life? .........................
2. Is the relationship mentioned above coincidence or true cause and effect? .................................
3. The experimenting that was done by Metchnikoff to verify his conclusion was .................................
4. Is Metchnikoff's conclusion sound? .........................
MODIFIED ESSAY TYPE AND SEMI-OBJECTIVE QUESTIONS

Answer the following questions briefly and as accurately as possible.

Respiration is an extremely important process that is carried on:—

Where? .................................................................

When? .................................................................

For what purpose? ................................................

Using what gaseous element? ...................................

Releasing what gaseous compound? ............................

DETECTING RELATIONSHIPS; below are pairs of terms. You are to find a very important or fundamental similarity between the two items, and an important difference. Your answer is to be limited to from four to six words in each space.

Pairs of items          Similarity          Difference

1.  Diffusion          a.                      b.  
    Osmosis

2.  natural immunity   a.                      b.  
    artificial immunity

    malaria
    bubonic plague

COMBINATIONS SO MANY IMPOSSIBLE TO MEMORIZE
Demands actual active thinking over of facts. The facts may be memorized but the way to use them cannot for the student can have no idea what pair of ideas are to be compared.
CHAPTER VII

FORMS OF QUESTIONS SUITABLE FOR TESTING THE ACHIEVEMENTS
OF "LEAST" IMPORTANCE.

This chapter deals with forms of questions suitable for testing in accordance with the objectives of General Science IV and V. These objectives are not sharply marked off from the preceding and comprise six of the ten objectives of General Science IV and V.

To provide materials for worthy use of leisure

This objective was ranked fifth with an average weighting of eight percent.

As the outset it appears that the measurement of achievement toward this goal is extremely difficult. First, the complete attainment of the objective is not possible in school, but only in out-of-school hours or in life after leaving school. This fact makes the testing of achievement extremely difficult and next to impossible. Were it not for certain evidence one would be tempted to say that it is impossible to test for achievement here. Any programme of testing this objective is based on the hope that interests aroused in school life will carry over into adult life to a moderately high degree. Thorndike long ago gave some idea that this is possible in his investigations into how well pupil interests carry over into later functional activity, reporting correlations between these two of .66 and .89. This measure of the transfer of interests is at least as great as the carry-over of many of the ordinary topics taken in high school courses. A test on the increase interests would appear to yield a fairly valid measure of achievement in any particular course. Interests do not develop entirely on their own, from within as it were, but are the results of wider experiences that

1. Taken from Hull's report in "Aptitude Testing" page 190.
come from education. There is development of interests and this development should be measured.

Another difficulty with this objective lies in the interpretation of the word "worthy". The difficulties concerning the interpretation of the word have been mentioned in a preceding chapter.

The chief purposes of testing the progress toward this objective seem to be to measure the pupil's realization of the wealth of interesting material, which can be of tremendous interest if pursued further, and in measuring the teacher's success in arousing these interests.

Very few of the usual question types are suitable for testing achievement of this objective. True-false, completion, multiple choice, deductions all fall short of measuring with any degree of reliability that which is worth while. Most of them are basically opposed, for they are based on the philosophy of giving the student only certain choices which he must act upon, and not on the conception of the complete freedom of choice which marks most avocations. Desires and urges may drive activity in almost any direction. Those activities which injure the person morally, physically, and intellectually we choose to call "wrong". They may be so in the light of an omniscient judge, but our basis for judging is more or less arbitrary. It seems safe enough to consider any response which does not contravene these standards as a "correct response".

Practicability and costs are other factors influencing in no small measure the choice of avocations. The amount of leisure is another serious problem, although it appears to us now that the future will bring more for the great majority of persons.

The essay type of question undoubtedly serves a good purpose here for it leaves the horizon unobstructed, only bidding that the student
look east, west, north, or south for solutions. The evaluating of statements made by the student may be difficult but the specifications should demand only that the student realize the possibilities for avocational interests, and that the hobby or interest be not anti-social or anti-personal. If the student makes out a good case for a hobby that the examiner had not thought possible or interesting the examiner must accept his response.

Some questions that attempt to test in this manner are here presented:

1. Suppose that you have sufficient spare time and just enough money to provide the needed materials to cover expenses, what hobbies or interests of a scientific nature would you follow?

Or recognizing economic limitations, this:

2. Suppose that you have sufficient spare time, but not very much money, what hobbies or interests of a scientific bent or based upon what you find interest in a science course would you follow?

The essay type question should be tried again but this time with criteria for measuring the responses. Here is a good place to use it. As said before, most of the older methods of measuring essay type responses failed to provide the examiner with any scale of values.

Another type of question that measures a student’s ability to see opportunities for interesting activities is to limit the field of science and to ask him what possibilities there are in that realm. Such questions as this must take a moderately wide sampling in order not to militate against a person who does not happen to be either interested or able in the field selected. A suggested question of this type follows:

What hobbies or interesting activities do you think you could develop in the following fields?

Electricity?

Heat?

Light?

Buoyancy and Archimedes’ Principle?
Chemicals that affect human beings?
Minerals and Metallurgy?
Study of rocks?
Clothes and textiles?
Microscopy?
Study of sea animals?
Trees and flowers?

Another method of arriving at some idea of success in the development of avocations has been to use check lists. The usual check list is one of the "interest" type, and usually not very satisfactory, for it requires customarily a checking of each item in which the person examined is interested. Almost immediately the examinee "catches on" to the purpose of the test, and from that moment the results are not very valid. Sometimes to reduce this error somewhat the examinee is asked to evaluate each one of the items on some such basis as "like ...................... indifferent, .................. dislke". This lead if carried further can produce more accurate results so that the personal error of deliberate bias in order to obtain a higher score can be reduced to a very low level. The first attempt of the investigator to extend this technique was to have pupils evaluate given items numerically according to the stated scale of values, as follows:

Below is a list of hobbies, interesting activities, and such that are based upon some phase of science work such as you have been doing or are now doing. Read each item carefully, then place the number of the statement listed that best explains your opinion, reaction, or attitude to the item.

Use a 3 to show that you have been following those activities for some time now (that is, before this course, and still continue to do so).

Use a 2 ------- beside those hobbies, interests, activities that you have developed during this course, or have revived as a result of this course.
Use a 1......beside those activities which you were interested in but discontinued before you took this course.

Use a 0......beside those hobbies, activities in which you have absolutely no interest, nor had an interest in the past.

Use a -1......beside those activities that you think a student should not follow at all.

Use a -2......beside those hobbies, interests, activities that you once followed but have turned against as a result of this course.

HOBBIES, INTERESTS, ACTIVITIES.
EVALUATION.
1. Reading a good book on building of great projects, tunnels, etc.
2. Chemistry experiments on your own
3. Making explosives, nitroglycerine, etc.
4. Keeping a science scrap-book
5. Collecting and pressing flowers
6. Repairing electrical systems, doorbells, etc.

SCORING: This form of test could give two measures according to how it is scored. If we wish to find the efficacy of the course in developing these interests and hobbies we certainly must consider the number of 3's and 2's together with the number of 0's and -1's and -2's. This scoring is suggested:

Frequency of 3's multiplied by 2
Frequency of 2's multiplied by 1
Summation for gross positive scores.
Frequency of 1's multiplied by 1
Frequency of -1 multiplied by 2
Summation of negative scores.

Net score——Positive gross minus negative gross scores.

If a net negative score were obtained it might mean that the course has been unprofitable, or that the teacher is disliked. Further, it is quite conceivable that a person giving this result may have only a few interests in the course yet may follow them very vigorously. For failing to bring out the great intensity of these interests the scheme is at fault and would need to be supplemented, but the test does appear as if it would
give a good measure of extent of interest development.

To obtain some idea of the extent of the student's interest the frequency of the 2's and 2's could be summed with the frequency of the 1's.

Further work on this idea was done in the hope of obtaining some graphical arrangement that might give a more interesting analysis of the situation, using as a guide the work that has been done to develop personality contour graphs and similar statistical figures. If one desired to find whether a student's hobbies and leisure time activities ran towards chemistry or towards biology, the topics could be arranged in these groups. Further, if one desired to find the extent of the student's interests, the headings could be arranged in columns and a profile view obtained of the course's results and the pupil's reactions. The list presented to the student must contain those items that may be deemed socially satisfactory. To introduce negative factors is to introduce another problem entirely, that of evaluating what is socially worth-while. This last function is done best by specially trained persons.

With the foregoing rearrangements made the final form is now presented.

DIRECTIONS.

In reading through this exercise you will see two sections. The first gives you certain statements that you will use in evaluating the list of items which follows to make up the second section. Each statement has a number so that you can use the number instead of copying the statement and thus save time. Read the statements very carefully, then judge each item on the basis of these statements. Place a check mark in the column of the value you give.

Use a 4 to represent those hobbies, interests, activities that you have started or developed during this course. Place the check in Column 4.

Use a 3 with a check mark in the proper column to represent those hobbies, activities which you have been following for some time and still do.
Use a 2 to represent those activities or interests that you think you might want to adopt.

Use a 1 with a check mark in the proper column to represent those hobbies, activities that you ceased to follow before you took this course.

Use a 0 check mark for hobbies, and activities in which you have absolutely no interest, or desire to follow, nor had in the past.

Use a -1 check mark for those activities which you think a student should not follow at all.

Use a -2 check mark to represent hobbies, activities, etc. that you once followed but have dropped or turned against as a direct result of this course.

Example:

<table>
<thead>
<tr>
<th>ITEM</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>-1</th>
<th>-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Growing prize-winning chrysanthemums</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Collecting diamonds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Studying the stars</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SCORING:** The same plan of scoring could be followed as for the last method with this advantage that the response sheet when complete forms a
SCORING: (cont.)
profile of interests in scientific fields. However the profile obtained depends upon the plan or arrangement of the test. If it is desired to see which unit promotes most interests the items of the test must be arranged on a parallel unit basis.

In connection with this modification of the check list idea the examiner prepared a list of one hundred twenty possible "general" type of avocations based on science and asked a class of Grade IX students, one of Grade X, and one of Grade XI (having taken General Science I, II, III and IV respectively the preceding year) to evaluate their avocational interests in science. The purpose of this test was to see if interests increase with the more science work taken. To gain some idea of the value of the measurement of the increase of interests one should remember Thorndike's research on the carry-over of adolescent interests into adult life where he found a correlation of .66.

The results of the experiment conducted by the experimenter seem to show that the "seven value" scale practically eliminates the "padding" evil of the usual check-lists wherein students often guess shrewdly in attempts to make high marks. The "like" column that is so frequently found in check-lists has been broken up and made more specific so that a vague liking must be clarified in the student's mind before he enters his score on the list. The results show a distinct increase of the:

1. Realization of possible avocations (truly a creation of interests) in each year of science.
2. Actual hobbies developed from year to year.

There is no dropping of interests due to dislike created by the science courses. Whether the special sciences create more interests than General Science could not be answered from this experiment, but the res-

1. See Experiment 1 of Appendix for further results.
Responses of several Grade XI students who had taken one year of special science work suggested the possibility of this in certain fields. The form of test seemingly does measure growth of interest and thus is valid for the purpose.
To Appreciate Achievements in the Field of Science, and the Contributions of Scientists to the Modern World

This objective was ranked sixth and given an average weighting of nearly eight per cent. The objective is rather straightforward with no implications that are not readily seen in the statement as it stands.

The appreciation of the contributions to our social heritage made by men of science can be measured moderately easily, with scarcely any modification of many of the existing question types. Those forms of questions which demand associations, comparisons, or evaluations seem to fit this objective better than those lacking this element. For this reason essay response questions comparing the results of two men's contributions or their effects on society are very good, but the difficulty of evaluating the responses is greater than for the objective questions. Matching questions are fairly good, for they seem to evaluate this appreciation to the extent of associating or of differentiating among those who made these certain contributions. Direct questions asking students to name the chief contribution or contributions of certain men are quite useful, but have the slight lack of complete objectivity when several possible responses could be counted as correct. With a properly prepared table of specifications for each set of questions the marking should become quite objective. Completion and multiple choice questions tend to emphasize more the factual recognition rather than the appreciation of the contributions.

In passing, it seems that one warning should not be overlooked. While it would be very fine to compare the results of scientists in bettering social conditions with those of warriors and statesmen, the objective type of question for this measuring of appreciation would seem to be unsatisfactory on a science paper because the student would soon "catch on" to the situation and would give his answers a bias towards the scientific
contributions. If the questions and comparisons are kept within the scientific field this objection no longer holds. Comparisons between the contribution of a scientist and another person outside of the field of science are best done in debates. Essay questions are moderately fair testing means, but the evaluation is extremely difficult and it is essential that a key be prepared.

TYPES OF QUESTIONS.

1. Essay type; Compare the contributions to humanity of the two Frenchmen Napoleon and Pasteur.

SCORING: The specifications demand an appreciation or evaluation of the contributions. The question does not inform us on what basis the student is to judge, and is to that extent poor. If it is on the utilitarian basis of the greatest good to the greatest number, this should be stated.

The question would be better if re-worded;
Compare the contributions to the welfare of humanity made by the two Frenchmen Napoleon and Pasteur. Judge on the bases of the extent that health, wealth, and happiness were increased or reduced.

Such a question now permits the preparation of a marking key.

In favor of each you should expect to find these points advanced:

Napoleon; caused the revision of weights and measure to be completed, and the metric system to be introduced. Stimulated the methods of preserving food, of course to make his army more independent of time and fortune, so that Appert's discovery of canning was the result. Emphasized the importance of diet for an army.

Pasteur; brought solutions to the difficulties in several of France's major sources of income; the cause and cure of anthrax; of the souring of wines; of the dying of the silkworms; a cure for hydrophobia; his research led the way to antiseptic surgery by Lister; and established the germ theory of disease. Swept away much of the ignorance and superstition in France. Has saved millions of lives, and continues to do so, though dead. All nations have profited.

Against each; Napoleon brought great disturbance to the world; much misery and many dead "heroes" and innocents; his projects unbalanced Europe for many years after his fall. Pasteur's life was exemplary, for we have no records of unkind or cruel actions on his part, nor have his discoveries decreased health, wealth or happiness.
Having prepared the specifications, we are then able to evaluate the points of credit, and thus help to make evaluation more objective.

2. Direct statement type, semi-objective;

What do you consider the chief contribution to society of each of the following scientists?

a. Lavoisier ................................................

b. Galileo ....................................................

c. Faraday ....................................................

b. Pasteur ....................................................

e. Edison .....................................................

f. de Forest ..................................................

g. Koch .......................................................  

h. Dalton .....................................................

i. Newton .....................................................

j. Burbank .....................................................

k. Saunders ...................................................

l. Watt .......................................................  

m. Baekeland ................................................

3. Multiple choice questions.


II. Whose discoveries have been developed to such an extent that vast amounts of moderately cheap energy hitherto wasted can now be put to work for man? a. Lavoisier, b. Mendel, c. Morse, d. Archimedes, e. Faraday, f. Volta, g. Ohm, h. Bell.

4. Recall or completion type;

Fill in the blank with the name of the person who made the following contribution to society.
a. insisted that before any germ be classed as the cause of a particular disease it must first be isolated, then grown, replanted to recause the disease.

b. really the first of modern scientists who placed more faith in what he saw when he experimented than in the statement of old and long dead authorities.

c. suggested the fruitful theory that only whole atoms can unite with only whole atoms in the usual chemical reactions.

d. in 1862 showed the value of vaccination against sheep anthrax on two groups of sheep.

A valuable type of question, and possibly the best, is to give a name of the person making a very simple scientific discovery and to ask the student to supply the modern applications of this. For instance:

Below is a list of very simple discoveries made by investigators in certain fields, in the blank name our chief modern applications of this idea:-

1. In the 1840's Swan made a piece of metal wire glow red hot when he passed electricity through it. Modern application ........................................

2. About 1700 Hauptman and Longius, with the aid of crude lenses, discovered little wiggling things in the pus from lesions of certain diseases. Modern application ........................................

3. Faraday accidentally moved a coil of wire through the field of a magnet to produce a temporary current in the wire. Modern application ........................................

4. Lavoisier found that the weight of oxygen and mercury produced by heating mercuric oxide equalled the weight of the mercuric oxide. Modern application ........................................

5. In 1883 Edison discovered that from a wire near another one heated red hot, and both in a vacuum, a negative current could be drawn. Modern application ........................................

6. When a tube containing air under reduced pressure of a partial vacuum has electricity passed through it a glow develops. Modern application ........................................

---

1. Suggested by Prof. C. H. Wood of the University of British Columbia.
This type of question is essentially a completion question and possesses a high degree of reliability in general. The correct responses are limited in number to a reasonable degree, for seldom will more than three or four major applications result.

The converse type of question is possible but it demands a greater familiarity with scientific history than is usually found in high school students.

For instance, to ask:

"What very simple discovery resulted in the development of our modern steam engines?" or "The modern steam engine has been developed from the simple . . . . . . . . of . . . . . . . . . . . . . . . . ." might produce these answers:

a. Watt's "invention" of the reciprocating steam engine.

b. Newcomen's "air pressure and steam" engine.

c. Hero's steam machine.

Likewise the lighting effect caused by electricity passing through small enough wires had been seen by many observers before Swan's work, which in turn led on to Edison's work.

If this type of question is used within a defined boundary, say the contents of certain text books it could be quite useful. It would be cast best in a completion form.

However, the first type of this pair could be used to go back to the prime discoveries with probably no increase in difficulty.

Another variation of the first is to name an invention or a discovery and ask how it has been applied, omitting all reference to the names of persons.

1. Also suggested by Prof. C. B. Wood.
To enable the Student to counteract superstition and to correct erroneous beliefs through the application of Scientific Principles.

This objective was ranked seventh and given a weighting of seven percent.

The objective at first glance would seem to duplicate much of objective number two, that is, the development of the ability to use the scientific method, for it states that its chief means of attaining its end is the application of the scientific method. A teacher might wonder if he limited his attention to the development of scientific method and thinking that it might be enough in itself, because there should be considerable transfer. Really the teacher must see to it that these common elements are brought into the consciousness of the pupils before the transfer will take place.

It would not be safe to rely only on training in the use of the scientific method because many wrong beliefs come to us without critical analysis. they creep into our stock of ideas when we are children, they insinuate themselves in news and in advertising. The less our training has been in the ways in which wrong beliefs are absorbed the harder it is for us to detect the presence of them. Training in critical analysis, logic, and the methods employed deliberately by advertisers and propagandists will assist materially in checking wrong beliefs, half-truths, and superstitions. Many people seem to possess compartments in their "mental stock" for they hold ideas in one field of thought that are philosophically contradictory to those held in another field.

Teachers are human and are likely to overlook teaching this aspect of science training unless the objective is stated clearly. In these days of turmoil amid fountains of propaganda this objective can serve a very valuable function in preserving democratic institutions from the parasitic
growths that tend to develop and eventually kill them. On this matter science really should join forces with Social Studies.

In view of the foregoing discussion the objective merits a place for what sub-objectives should a teacher strive?

This is a necessary question both from the point of view of teaching and from testing. In high school work the training of students should be toward the counteracting of some of the more insidious wrong beliefs than against the simpler, common superstitions such as those concerning ground-hog day, walking under ladders, and breaking mirrors. The teacher must oppose these, of course, with the reasons and truth of each situation because there is a continual reappearance of these beliefs, although they are slowly losing place in the minds of the people at large.

In high school the teacher should set himself more against wrong beliefs, the "ostrich attitude", fallacies of thought that are or may be widespread or peculiarly local, half-truths, plausibilities without a true foundation, astrology, fortune telling, propaganda, and the ever-increasing menace of ruthless, yet brilliantly skilful advertising carried on by certain types of business, the enervating effect of repetition ("Repetition makes for truth".), "authority", and the almost slavish veneration for the words "science" and "scientific" which have been prostituted by person's with ulterior purposes.

A pause here to analyze some of the techniques of propagandists would not be amiss as it will help to clarify the objective. To achieve their ends propagandists use these techniques in the main. To oppose these might be considered as some of the sub-objectives.
1. "Circumlocutions to reduce chances of ready analysis of the situation.

2. "Red herrings"; drawing some other issue or idea across the path of the thoughts which they had started to follow in order to lead the minds of the hearers or readers from the real issue or condition.

3. "Labelling"; damning something by giving it an opprobrious name such as "worn-out theories", "antiquated", "fad", "frill"; or oppositely bolstering some weak case by the use of favorable or strong labels, such as "scientific", "accepted custom", "new", "old and tried"; almost any term can be used either way, showing our enslavement to words and lack of understanding of ideas.

4. "Climbing on the band-wagon", "everybody's doing it". These are the stock-in-trade methods of party healers and workers, and nation-wide advertising of cigarettes and such. Sheep-like "follow-the-leader" is the game to reduce thinking to a minimum. Mental integrity often gives way to social convenience.

5. "Authority", and the modern form in advertising "endorsement". "So-and-so says so", is the ultimate of too many arguments. We had enough of this for two thousand years after Aristotle. To quote an example from another field:— If Jean Astruc's authority in France and at the court of Louis xiv had not been so great the germ theory of disease advanced by such men as Hauptmann, Langius, Kircher, and Saguens would have developed into the useful state almost two hundred years before it did and would have prevented untold suffering. Jean Astruc's analysis of the theory showed that he had a clear understanding of the implications of this theory, yet he preferred not to throw over the entire theory of disease then held. With brilliant rhetorical "red herrings", circumlocutions, and reference to the prestige of the leaders in his profession he strangled the infant born of the new discovery of the microscope.

"Endorsements" are modern advertising forms of this same weakness of humans. Endorsements, usually for a price, are found in almost every large advertisement in magazines, and are extremely difficult to pick out from the truthful statements.

6. "Repetition". Repeat a wrong belief enough times and it tends to simulate truth. Repetition is an opiate.

7. Another error, mainly in reasoning, is the ancient "Post hoc, ergo propter hoc". This has been with us like the plague in days gone by, and can be eliminated by the same methods, the application of the sanitation of the scientific method. This certainly should be one of the items for which we should teach and test, for this is the basis of many of our superstitions which are based usually on coincidence.

In order to test for achievement of the objective almost all the common types of questions can be used. One of the most promising is the "deduction from premises or statements" because errors and fallacies
can be incorporated into the question without making them too obvious. In life most of our wrong beliefs insinuate their way into our mental stock, together with the true beliefs. The incorporation of an error or fallacy into a very plausible paragraph mixed with true statements is a very life-like situation. Moreover, the ability to deal with longer and more difficult problems than in the junior high and elementary grades should be developed in the high school, because life situations often are very complex. The very brief completion tests, multiple choice, and true-false questions are very useful to check the more common errors and superstitions with which the student by the time he reaches high school should be quite familiar. True-false questions serve a very useful purpose as means to make rapid inventories of the stock of fallacious beliefs held by a person. All these types are useful, but testing in high school should go a bit further than memorizer acquaintance with these wrong beliefs. That is why the "deduction from statements" type of question should be used more, for it does call forth a higher degree of mental activity than memorizer type questions. It is a more life-like problem insofar as it can present an old error in a new guise, and is very much the way these problems are met in day to day experiences: To be able to detect these fallacies when the first acquaintance is made with a situation is akin to aseptic surgery and hygiene; "An ounce of prevention is worth a pound of cure". When the infection becomes deeper it is much harder to eliminate. Frequency of meeting it often dulls resistance to it; as Pope says in his "Essay on Man"

"Vice is a monster of so frightful mien,
As to be hated needs but to be seen;
Yet seen too oft, familiar with her face,
We first endure, then pity, then embrace."
An example of the "deduction from statement" question is given here:

DIRECTIONS: Read the paragraph carefully then answer the questions below. Each statement that you find in the list after the paragraph is to be judged by you as

- "True" for which you are to place a plus sign (+) in the parentheses in front of it.
- "Unsound or false judgment"; place a minus sign (−) in the parentheses.
- "Having no relation to the problem"; then place a zero (0) in the parentheses.

"An investigator, operating under the authority of the American Medical Association, found that, in examining the products and advertising of four hair-dye manufacturers, each one contained at least one of these poisons in the preparation; lead acetate, silver nitrate, copper sulphate, paraphenylendiamin. One of the manufacturers strongly maintained in his advertising that his products contained no silver nitrate, the dangerous paraphenylendiamin, or copper sulphate; another's advertisement claimed that no poisonous lead acetate, copper sulphate, or paraphenylendiamin, was used; the third produced a doctor's statement that no harmful lead acetate, silver nitrate, or paraphenylendiamin enters the formula; the fourth one claimed that his product was free of paraphenylendiamin. All advertisers used many testimonials from customers testifying to the worth of the products." (The investigator was well trained.)

( ) 1. The third hair dye contained the poison paraphenylendiamin.
( ) 2. The investigator is only one man against the many who wrote testimonials so his word should not be accepted.
( ) 3. The first hair dye probably contained lead acetate.
( ) 4. The purchaser deserves to be "stung" if she is not sharp enough to find out things for herself.
( ) 5. Germany has banned all hair dyes containing chromium, cadmium, and lead salts.
( ) 6. The advertisements correctly advertised the products for no manufacturer would put harmful substances into a preparation.
( ) 7. All these manufacturers followed the practice of telling half-truths in order to hide the real truth.
( ) 8. Many women use these dyes so that the dyes must be quite satisfactory.
( ) 9. The investigator was a highly qualified thorough man so that his decisions weigh more than all the testimonials, many of which might have been forged.
( ) 10. In order to cover their own shortcomings the advertisers drew attention to the harmful ingredients of other hair dyes.

SCORING: Because there are three possible values to be given each statement the questions more nearly approximate multiple choice than true-false. In three-answer questions to subtract one-half the wrong responses from the sum of the correct ones is the usual procedure.
TRUE-FALSE QUESTIONS.

These questions should not be the usual very brief question that they so frequently are, for very seldom do we find our errors and wrong beliefs eroded of plausibilities. Four questions have been suggested of the true-false type which can be used very effectively here for the purpose of testing ability to detect fallacies.

DIRECTIONS: Mark each statement "R" for right or "W" for wrong according to your best judgment.

1. A small vial of mercury worn on a pendant around the neck of a person going on a long sea journey will prevent that person from becoming sea-sick.

2. It is quite possible to travel to the moon in a dirigible or balloon if only scientists could discover a way to drive the craft out beyond the region of strong gravitational pull of our earth.

3. There is a strong possibility that forces are still unknown to us, just as electricity was practically unknown over 150 years ago.

4. Engineers are making steady progress reducing friction losses in machines. It will be possible one day to develop engines and motors with our fine workmanship that once set going will continue to do so until we wish them to stop.

(To show the extent to which wrong beliefs spread even in educated persons may the investigator state that the first one is a belief that he met as held by two lady teachers in Vancouver who were telling of their plans for a trip to England and France. They had asked a druggist in town for some medicine or cure for sea-sickness. He had advised that mercury in a small vial to be worn about the neck as a pendant as being one of the most frequently used methods of warding off sea-sickness. The teachers asked for a vial each. Each vial of mercury was well sealed. On their return from the trip one was asked how she got along on the ocean voyage and she remarked that she had been quite sea-sick but guessed that it must have been from the fact that she kept the vial in her handbag as she did not like the appearance of the pendant when hung around her neck. She presented the investigator with the vial of mercury which was kept for many years. With the other teacher he was unable to get in touch.)

If the last question were reduced to a blank statement about perpetual motion there would scarcely be one student who would not give a correct response, but as the question stands many who do not recognize some stock phrase will succumb to the plausibility.
The above form of question could be improved still further by asking the students to correct all statements which they do not think to be correct.
This objective was ranked eighth and given a weighting of six percent. It would imply that the student should realize these factors at least when he decides to follow a particular vocation:

The Student or Apprentice

1. his own ability in the particular field
2. his interests in the type of work done. This is not identical with #1, although it doubtless contributes to achievement.
3. his health and physical abilities.

The Vocation or Job

4. the training or knowledge demanded by the vocation.
5. the possibilities and opportunities of the vocation.
6. the monetary returns from the vocation.
7. the health conditions or hazards.

From his regular work in science the student should get a fair idea of his capabilities, particularly from the results of testing for the first three objectives. His interest can be determined mainly by himself; because it is a personal matter. From his courses, and from discussions that should take place from time to time in his class and among his friends, he should obtain an inkling of what vocations there are in which science training is necessary, but this information would not be sufficient and must be supplemented mainly through the guidance courses. From the science work he should gain a knowledge of the laws, principles, facts, that any particular profession or vocation based on science demands. The monetary returns and the health conditions should be studied mainly in the guidance group and the health classes respectively.
main function of the testing programme as it pertains to this objective would deal with factor four mentioned above, and slightly with the others because these others will be tested in other parts of the course or in other courses.

A modification of the multiple response question type would serve very well for measuring knowledge concerning the techniques demanded by any particular vocation that is based on science.

two samples of this type are offered below:

DIRECTIONS: Below you will find a statement with a list of items of knowledge, techniques, and activities that may be necessary for a person to possess who intends to enter the vocations mentioned; place a check mark in the blank in front of each item that you think the person should possess.

1. A druggist properly trained in pharmacy, and not just a salesman in a drug store, should have considerable ability and knowledge concerning

   ( ) a. the measuring of electricity
   ( ) b. soil conditions and erosion.
   ( ) c. weighing chemicals
   ( ) d. growing plants and caring for them.
   ( ) e. the grafting of trees and shrubs
   ( ) f. identifying chemicals, compounds
   ( ) g. atmospheric changes
   ( ) h. the effect of age on organic compounds.
   ( ) i. spectrum colors and light theories.

2. A girl contemplating becoming a nurse, besides possessing the attributes of good health, patience, strength, and cheerfulness should know considerable about these facts, principles, and techniques.

   ( ) a. enzyme action and digestion
   ( ) b. the laws of gravitation
   ( ) c. various rocks and minerals
   ( ) d. communicable diseases.
   ( ) e. ventilation
   ( ) f. refrigeration systems
   ( ) g. wonderful new medicines widely advertised.
   ( ) h. germ theory of disease.
   ( ) i. the nature of heat
   ( ) j. how mammals reproduce
   ( ) k. breathing system of mammals
   ( ) l. dynamos and electric motors
   ( ) m. Lenz's Laws.
   ( ) n. to every action there is an equal and opposite reaction.
   ( ) o. internal combustion engines.
   ( ) p. preparing food and diets.
   ( ) q. six simple machines.
   ( ) r. x-rays.
   ( ) s. osmosis and diffusion
   ( ) t. gradation and diastrophism
   ( ) u. streamlining and transport.
-SCORING: When no set numbers of responses are demanded from the pupil it seems fairer to count all his correct responses. These should be interpreted to mean all the check marks in the correct places and those unchecked that should be unchecked. From this sum the errors should be subtracted as in true-false types, of which type question this is really a form. If the subtraction is not done a person could score full value by checking every item, and the test would be useless to measure his knowledge of what would be needed in the way of training. If, on the other hand, the question is worded this way, "Check the four (or five, six, etc.) items of the list which you think a person should know in order to enter this profession (or vocation)," then the question must be reorganized. In recasting the question care must be exercised to include just the four correct responses with a goodly number of "off-color" and incorrect ones.

Another type of test item is what might be called a case study. This could be the record of some person, and the decision that he had to make in the matter of choosing a vocation. The person's marks in sciences could be examined and the student asked to make a judgment as to what line of endeavor he would advise a person with such a record to follow. Questions could be asked also about other factors affecting such a decision to see if the student realizes the importance of factors other than those of an academic nature. An example follows:

"Elfreda Garrett is not quite sure whether to become a nurse or not. She must make up her mind this year because the hospital where she wishes to train demands of its student nurses that they have high school work in chemistry, Biology, and Home Economics (cooking, foods, dietetics). She will enter her last year at high school next year and she can work in these subjects if she desires. Below is her record;
The required home economics course she has taken already with these marks for the four quarterly rankings; B, C, B, B.

Her marks in General Science were:

### First Year

<table>
<thead>
<tr>
<th>Unit</th>
<th>General Science Topics</th>
<th>Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit I</td>
<td>(matter, atoms, molecules, formulae, chemical laws)</td>
<td>C</td>
</tr>
<tr>
<td>Unit II</td>
<td>(energy, force, power, machines, electricity, magnetism)</td>
<td>D</td>
</tr>
<tr>
<td>Unit III</td>
<td>(atmosphere, gases, climate weather)</td>
<td>C</td>
</tr>
<tr>
<td>Unit IV</td>
<td>(water, its composition, pressure, purification, life)</td>
<td>D</td>
</tr>
<tr>
<td>Unit V</td>
<td>(plants, use and functions of parts, and protection)</td>
<td>D</td>
</tr>
</tbody>
</table>

### Second Year

<table>
<thead>
<tr>
<th>Unit</th>
<th>General Science Topics</th>
<th>Mark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit I</td>
<td>(life processes common to all living things; adaptation to environment; heredity; reproduction)</td>
<td>C</td>
</tr>
<tr>
<td>Unit II</td>
<td>(transportation, role of electricity, internal combustion engines, safety measure)</td>
<td>D</td>
</tr>
<tr>
<td>Unit III</td>
<td>(radiant energy, communication, light, sound, radio)</td>
<td>D</td>
</tr>
<tr>
<td>Unit IV</td>
<td>(carbon and nitrogen cycles; chemical properties of these compounds of these elements)</td>
<td>C</td>
</tr>
<tr>
<td>Unit V</td>
<td>(Earth's crust and minerals; methods of obtaining and refining needed minerals; carbon, sulphur, salts)</td>
<td>E</td>
</tr>
</tbody>
</table>

Place a check mark in the blank opposite the best answer according to your opinion in each of these three questions. For the last one you are to do more. You must check the reasons or statements below that you would use to support or defend your decision in question 3.

1. Judging on the basis of her General Science marks, next year in Chemistry Elfreda would
   - ( ) a. do extremely well
   - ( ) b. do well
   - ( ) c. "get by" with a bit of difficulty
   - ( ) d. learn very little
   - ( ) e. fail completely.

2. Judging likewise from her record in General Science, Biology for Elfreda would
   - ( ) a. be extremely difficult
   - ( ) b. be successfully handled
   - ( ) c. cause her to fail next year
   - ( ) d. not be of much use
   - ( ) e. be unnecessary.

3. Elfreda should
   - ( ) a. become a nurse without any more hesitation
   - ( ) b. not consider nursing any further
   - ( ) c. consider many other factors as well

### REASONS:

1. Biology offers no serious obstacle to her progress.
2. Chemistry will give too much trouble to be worth it.
3. She must like "looking after people".
4. She could pass in chemistry with much hard work.
5. The Chemistry course will be much easier next year.
6. They say that the Biology course is harder next year.
7. Her record in Home Economics is not good enough.
8. Strength and health are important factors.
9. Patience is a necessity in the vocation of a nurse.
10. Her Home Economics record is quite favorable.

Aptitude tests might be given as in Stanford University, but this is very much out of the field for the science department. Again, many of the factors tested by the Stanford Scientific Aptitude Test are tested in a good testing programme that covers General Science IV and V. Any test that measures achievement in the use of the scientific method will cover most of the work of testing that the Stanford Scientific Aptitude Test covers. Further testing than this is the duty of some other department, probably that of Guidance. Likewise, tests of vocational interests lie in the domain of the Guidance department, for any testing programme that would be worth while would be too extensive to be included in the science courses.

That interests are a very potent factor in the choice of a vocation, even more important than ability, is shown by Franklin's research, an abstract being presented herewith:

<table>
<thead>
<tr>
<th>Grade VII B Pupils</th>
<th>Number</th>
<th>Average I.Q.</th>
<th>Av. Score on Clerical Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average non-clerical</td>
<td>500</td>
<td>106.5</td>
<td>34.8</td>
</tr>
<tr>
<td>Experimental group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desiring clerical career</td>
<td>135</td>
<td>99.3</td>
<td>47.5</td>
</tr>
</tbody>
</table>

Two standard group intelligence tests were given; the clerical test was from the Columbia Institute of Educational Research. Additional evidence of the value of interest in vocational guidance can be found in Thorndike's report.

1. Adapted from Hull's report in Aptitude Testing, page 191 (G.T. Hull)
He found that in general the order of a person's interests for any given period correlated .89 with the order of the same person's estimates of his abilities for the corresponding period. Bridges and Dollinger found student interests correlated with student estimates of abilities to a degree of .57. It would seem from all this that interest is a valuable criterion to use in deciding vocational matters. It probably follows that a very fruitful way to achieve vocational guidance in science would be to develop interests in science, and to give a sense of success in the field. At present this degree of success is rigid, fixed by the theory that marks or grades must be distributed according to a normal curve of distribution. We are bound to rank 25% very low, 50% average, and 25% as very successful. It does not offer very much encouragement for the weaker students, who, no matter how hard they work and what they actually may accomplish or know, are always struggling against the competition of abler students. Truly are our testing procedures extremely important and influential. It seems necessary to adopt some system of reporting progress that encourages more pupils and develops their interests.

In connection with the choice of a vocation the interest check list developed under the worthy use of leisure could be of some prognostic value.

The carry-over of interest may be assumed to be approximately similar to interest increase in each course. Direct evidence needs to be collected from interests functioning in later life in order to correlate with prognostic values.

3. See Appendix, Experiment 1, pp. IV-XIII
To develop the ability to perform simple experiments and thus to appreciate the scientific basis of science.

This objective was ranked ninth and received a weighting of seven percent. In view of the fact that it overlaps other objectives in part, the apportionment of marks or credits will need be somewhat arbitrary. Because of its overlapping the second objective dealing with the development of scientific thinking it is possibly better to allocate the major part of the credits or marks to achievements of a nature similar to manipulation, accomplishing general laboratory techniques, care and cleanliness, and planning an attack upon a problem. Any ability in experimenting would seem to be a composite of these minor abilities above and also of the ability to organize data and to deduce valid conclusions from the data obtained. The latter two minor objectives are part of the scientific method which is mentioned in the second objective. On the other hand manipulation and organizing plans bear a very close relationship to resourcefulness and adaptability. With these two boundaries in mind it would seem safe for the teacher and examiner to interpret the objective as throwing more emphasis upon the actual doing of experimental work and not upon what is called commonly "book-knowledge". It is true "learning by doing". The simplest of experimental problems should be attempted for in these the essences the experiment is found. Doing many simple experiments will produce greater mastery of the method of experimentation by habituating the student than will the attempting of a few long, involved experiments.

No one can develop a great ability in experimentation merely by watching others do the experiments, or by reading about them. Granted, an observer can learn something by so doing; if he is well-experienced he may learn much by observing and comparing with his own technique. The only
real measure of achievement of this goal is to measure the achievement directly by having the student perform the experiments, or manipulations, or techniques. For these experiments or techniques to be used as test items the teacher should have clearly in mind the purpose of each, and should work out a scale of values, just as for other examinations.

Indirect evidence bearing upon this aspect of the problem developed from experiments conducted by the writer. Correlation between "pencil-and-paper" tests and actual manipulations were low in several of the experiments and never higher than .77.

The testing of achievement by practical work in using experimental techniques introduces many difficulties, chief among which are the needs of equipment and demands for space. Because experimentation demands room it is not as convenient a method of testing as paper and pencil tests are. Again, the student is up on his feet and his eyes can not be limited in range, whether he is deliberately trying to find outside assistance or not. Small booths or arrangements similar to these are practically a necessity for this type of examination when the size of the class exceeds even fifteen students. Most classes in General Science in high school run between thirty-five and forty-five. In the larger classes practical examination of these abilities is very difficult owing to the lack of both space and privacy in testing. Only the simplest of techniques could be measured in this way then, for one teacher could not hope to check and credit forty doing a "job" that takes on the average about four to five minutes. It is physically impossible if he must supervise to prevent the spread of information and assistance. Practically no schools have equipment in the way of booths that could be used. This may be one of the reasons that so little has been done along this line to measure 1. See Appendix Experiments 11, 111, LV pp.XIV to XXVI.
practical achievement. For these many reasons teachers in science desire and try to develop paper tests of these achievements that have a sufficiently high correlation with the actual to make them worthwhile.

For the sub-objectives of planning the procedure, and for making valid conclusions the pencil and paper test is equal to any other means. When actual manipulation is involved, or when one must go to the prime sources for his data, pencil and paper tests no longer seem to suffice.

Two chief means of measuring achievement in actual experimenting are available. One is to keep a record of the student’s experimentation each period that he is expected to work in that way. The other is the "practical" or performance examination mentioned above. Practical examinations are not at all common in high schools, but they are used more frequently in university work in several of the sciences. Sources of questions possible to use in "practical" or performance examinations would fall into these classifications:

1. Repetition of some experiment done in class work.
2. Modifications of experiments done in class work.
3. Entirely new experiments but not above the level of the grade ability.
4. Certain small techniques, "tricks", and so forth that might be called the building stones of the experimental methods.

The Persing Laboratory Test in Chemistry has attempted to test for the last, and sets questions that demand the recall of information that would fall in the first group. No literature has been found to show how valid this test is for this particular purpose, as the authors make no mention of comparing the test results with actual manipulation and test
questions. Much of the test is actually factual material that could be measured equally well without diagrams. A certain amount of guidance concerning the validity of diagram tests as against the actual manipulation of things can be obtained from some of the work of Dr. H. A. Toops.

Moreover, research conducted by Dr. H. A. Ruger in 1922 would suggest strongly that the ability to manipulate things and ability to handle ideas (intelligence) are not likely the same abilities. In checking the correlation between ability to manipulate things on the Stenquist Assembly Test with general intelligence correlations were quite low, ranging from .06 to .11 for men and women of a Summer Session group to .41 and .15 respectively for a Winter Session group. It is quite possible, for a person to rank a genius on one and an idiot on the other.

It is advisable to test directly when measuring ability to manipulate things, to set up experiments, to perform experiments.

Toops reports that the correlation of pupil's scores on three paper tests using pictures compared with success in shop work, etc., was lower than the correlations of the scores on the Stenquist Assembly test with shop success. Picture tests correlate fairly highly with intelligence, +.60 but the Stenquist scores correlated only +.42 (approximately) with intelligence. General Trade and Mechanical Interest Tests correlate low with Army Alpha.

Some examples of questions in practical work and experimental procedures involving actual manipulation will be presented here:

DIRECTIONS: You are being tested on your ability to handle and arrange things necessary in doing experiments:

1. Toops "Tests for Vocational Guidance of Children Thirteen to Sixteen," Cont. to Education #136 Teachers College Columbia pp. 24-25
2. Toops "Tests for Vocational Guidance of Children Thirteen to Sixteen," Cont. to Education #136 Teachers College Columbia pp. 24-25
You are to pour some acid from the bottle into the test tube so that the instructor can watch you carefully. Pour about 5 c.c. and then return it to the bottle after the signal to do so has been given.

Suggested scoring:
4 for a completely successful performance covering: the holds of the stopper in the usually accepted manner between the fingers of either hand so that the plug part is away from the articles handled; the clearance of clothing, etc.; tilts bottle carefully and pours gently; replaces stopper, label side of the bottle in palm towards palm of hand, or up.

3 for any procedure that differs slightly on any but the last item.
2 for a performance that is rather careless but does not spill any acid, or a good performance up to here but fails to return stopper when return of acid is made.
1 for a poor performance without return of stopper.
0 for any performance which spills acid, providing that it is cleared away.
-1 for any performance which after spilling the acid fails to wipe it up with rags or dispose of it satisfactorily.

Weigh the stone which you are given as accurately as you can. Record the number of the stone, and also its weight. Do likewise with the piece of wood.

SCORING: Here the scoring is very easily made, for the only point that the instructor must decide is what tolerance will be permitted. This, of course, depends in large measure upon equipment.
Two points for each satisfactory weighing.
One point might be allowed for weights slightly off, say by one-fifth or one-quarter gram.
Zero for all other weights.
Instructor must be prepared for change in weight of samples due to breakage or other injury.

On the demonstration table is a series of containers with liquids in each. You are to record the letter of each one and after it place your measurement of the volume of liquid in that container.

SCORING:
One point each.

You are to arrange the apparatus allotted to you so that it could be used to prepare and collect quantities of a gas that will not dissolve appreciably in water.
SCORING:
Five points for a completely set up arrangement; deduct one mark for each error such as thistle tube not low enough (if used); delivery tube down into the reactants; loose cork or stopper; pneumatic trough not included; delivery tube not submerged; collecting bottles not filled with water.

V. Pin the printed label on the appropriate organ of the specimens provided.
(Specimens do not need to be named nor labels arranged into groups; preferably they should be jumbled so that identification would show more active thought.)
cortex, foot, sepal, mouth, pith, xylem, sepal, heart, anther, ovary, stomach, style, bract, phloem, adductor muscles, cambium, gills, stigma, mantle, petal, retractor muscle, medullary ray, siphon, filament.

SCORING:
One mark each correct labelling; or sum can be divided by two if weighting appears too heavy, etc.

All these questions can be turned into paper and pencil tests with various modifications. It is debatable whether this is a wise procedure or not. There has been an attempt to keep to the same subject matter for each question in turning the preceding into pencil and paper tests.

1. Explain briefly what you would do in pouring acid from the acid bottle into a test tube. If you cannot explain briefly in words a sketch may help you.

a. With the acid bottle stopper ..............................................

b. Grasping the bottle of acid ..............................................

c. Pouring liquid ..............................................................

d. With your body ..............................................................

e. When you have spilled some ..............................................

11. What is the weight in grams of a stone 11 c.c. in volume which, when placed on the left pan of the balance, needs a twenty gram weight, a 5 gram weight on the right pan, and has the beam rider or marker with its right edge at the subdivision 7 and its left edge at the first subdivision between the numbered marked points of 3 and 4? ..............
III. What is the volume, in each of the diagrammatic representations, of the liquid in each measuring vessel?

Water

Alcohol

Sea Water

Mercury


IV. Arrange the apparatus sketches so that when assembled the sketch that you make will show a complete set-up of apparatus suitable for preparing and collecting a gas like oxygen or hydrogen.
V. PLACE THE PROPER TERM AS LABEL IN THE BLANK SPACES PROVIDED.
cortex, foot, sepal, mouth, pith, xylem, sepal, heart, anther, ovary, stomach, style, bract, phloem, adductor muscles, cambium, gills, stigma, mantle, petal, retractor muscle, medullary ray, siphons, filament.

Figure XX

Figure XXI
SCORING: For these exercises the same scoring method could be applied as for the practical or experimental tests.

Other types of test items can be found on the Persing Laboratory test in Chemistry.

Although it has been out of fashion for many years to hold these views I think that a well worded essay type question can do much in finding whether a student can plan or organize his attack on a problem in a satisfactory manner. Some such question as this would make a student think and plan well.

"Given a pulley set, and weights, string, and scales if needed show how you would prove that in any machine "work in" is equal to "work out".

The scoring would have to be worked out on the basis of so much for plan, a certain value for his procedure in the actual manipulation, more for his recording and arranging data, and the final part score that for conclusion.
As a result of experimentation done by the investigator it was found that in the biological aspects most students who study the course by making use of specimens, diagrams and tests make slightly higher marks on the practical labelling tests. (This does not apply to dissection ability which was not tested.)

Judging from the results of experiments II, III, IV found in the Appendix the ability to perform simple experiments must be tested directly in actual experimental conditions and cannot be tested reliably by a "paper test" because the correlations though positive are too low, ranging from +.21 to +.77.

In later work carried on by the writer on the use and value of diagrams in testing, several improvements were made in the function of the tests. The preceding diagram tests emphasize names, or alternatively, the anatomical or structural knowledge. The functional or physiological aspects can be tested equally readily by diagrams simply by a revision of wording in accordance with this point of view. For instance, by referring to Figures XX, XXI, XXII it will be seen that suitable questions can be cast in either of these two ways:

a. By using the number opposite an organ or structure state:

1. The muscles which close the shells
2. The organ that obtains oxygen
3. The organ that does the burrowing in the sand
4. The organ that produces pollen
5. The tissue of the stem possessing the power to divide

b. State briefly the function of the following structures:

1. The organ numbered 9
2. The organ numbered 13
3. The organ numbered 19
4. The organ numbered 17

This form of question emphasizes function and demands considerable scientific knowledge, with a minimum of technological jargon.

1. See Appendix pp. XLI to XXVI.
To Develop the Desire to Read Scientific Literature.

This objective has been ranked tenth with a weighting of only five per cent. It definitely was considered the least important of all the ten objectives by the seventy-eight teachers and administrators who replied to the questionnaire (before September 15, 1938). The question might be raised as to whether or not this should be an objective of these science courses. The Science Revision Committee after lengthy reading and discussion of the objectives of these science courses decided that this is an appropriate one. Supporting evidence comes from the increasing frequency with which popular articles of a scientific flavour are included in the general reading matter of newspapers and magazines. To increase the desire of students to read material like this is certainly a worthy aim. It may or may not be difficult to accomplish this. To develop the desire to read is one thing. Enthusiasm in the teacher does this quite well. To measure this increased desire is very difficult.

As the objective stands it is admittedly hard to test achievement of it, for it directs the effort of the teacher along the line of developing a desire to read scientific literature and not directly to develop the ability to read it. Another point worth noting is that there are many levels of intricacy and difficulty of scientific literature. The objective has been criticized on the grounds that there is scientific literature of such weight that the specialists in the field have difficulty in understanding it. This is quite true, but more is implied by the Committee than the reading of scientific literature of a difficulty commensurate with the training in science that the student possesses at that time, or very slightly more difficult in order to permit his growth in understanding of science. "Scientific literature" should not be taken here to mean the
original sources which are usually replete with technical terms, but rather the good resumés of such works, abstracts, interpretations, semi-popular and sound popular articles in books and magazines.

Any test for this objective should test directly the desire to read such books and articles. The direct measure would be to give the student credit for the extent of improvement of this desire, but how can it be measured? By taking the student's own word directly? Or by trying to get some indirect evidence that would be more reliable than his professed achievement? Nearly all teachers would prefer the second method when they must deal with great numbers of students whom they cannot know very well. One indirect measure and a fairly valid one is to give the student credit for what books he has read during the course. The evaluating of this credit leads into a real difficulty. Here the teacher must be a bit arbitrary in his evaluating, for no way to measure and weight this achievement objectively has been developed. If this objective were given a weighting of 5% of the course value it would seem that the reading of two books of merit equal to Paul de Kruif's "Microbe Hunters" would satisfy, and more, the demands of this objective. Scarcely any person can read such books without obtaining a very vivid picture or impression of that field and its scientific procedures. So also with Slosson's "Creative Chemistry", and many others. The testing of the reading of such books can be done quite well by asking the student to answer a sheet of questions prepared on the book in question dealing with only the major points of the book, and the person's opinion of it. A teacher needs only to know that the person has obtained a fairly good idea of the book as a whole. The student should not be expected to remember everything as in a text-book. These test papers made by students should never be returned but destroyed lest other students less conscientious obtain them and memorize the test without
actually reading the book. The test items should be in forms suitable to
the material. This method is satisfactory for testing knowledge of the
books in the school library. It cannot be carried far beyond this step,
however.

The writer has come upon one method of evaluating the desire of
students to read scientific literature. This method gives a view of the
success of a method rather than the amount read by each student, although
the method could be revised to measure students individually. The method
was discovered mainly by accident. In preceding years the students in the
writer's biology classes were left free to read whatever was attractive
in the way of books with a biological scope. Out of sixty-six pupils of
the preceding two years only eight had read special books in biology to
the total of fourteen books. (One had read three, four two, and three
only one.) This represents about one-fifth (.213) of a book on scientific
matters per student.

This year the writer, being dissatisfied with this low reading
quota, assigned two books to be read by the students for credit. These
books were to be selected from two lists prepared, one book from each
list. Books were placed either in the group of biology applied to health
(such as "The Microbe Hunters," "Man vs. Microbes," "The Life of Louis
Pasteur"), or in the group of general biological interest (books by Ernest
Thompson Seton, Wm. Beebe, natural histories, books on mammals, birds,
insects, and similar ones).

During the school year until the end of March thirty-two students of
a class of thirty-five had read sixty-two books or an average of 1.94
books each or 1.77 for the entire class. Of course, this is required
reading but its concomitant seems to be a development of true interest
in this type of literature because nineteen additional books beyond the
prescribed ones have been read by the class for an average extra work of
.54 of a book per student. This compares very favorably with the .213
of free undirected reading of the preceding years. To assign a wise
minimum of books seems to develop the desire to read more literature of
a scientific nature.

To develop the desire to read one must develop the ability to read
also. At high school level the ability to read is not usually dependent
upon the training of eye movements, for these customarily have been de­
veloped in the elementary school. Only a few remedial cases of this nature
are likely to occur in high school. Rather should the training be along
different lines. Some of the main factors at this level in the develop­
ing of the understanding of scientific literature are the development of
vocabulary and the appreciation of the precision of technical terms, the
analysis of phrases and clauses to trace the exact relationships, and the
 provision of a good general background of knowledge or information. With
the first and third the science courses are directly concerned, but the
second properly falls in the domain of the English courses in high schools.

Testing for the increase in vocabulary is done already well enough in
most science courses in the placing of terms on ordinary achievement tests.
it is to be expected that the background of knowledge would be obtained
according to the ability and interest of the student. The amount of
testing of terms on any test should not exceed five percent of any
achievement test.

A fuller programme of testing this achievement would cover is here­
with suggested. The objective will have been satisfied when evidence
has been obtained by the teacher of:-
a. The amount read during the course.
b. the development of vocabulary necessary to understand reading matter.
c. the knowledge of sources of information according to the facilities of the school, or local library.
d. a measure of reading ability such as made by the Van Wagenen Reading Scales in biology.

One procedure dealing with "a" above has been advanced. As to the next item, "b" it is considered to be well enough done, and a familiar enough procedure by virtue of tests given for terms. Reading ability is obtainable in a way similar to the van Wagenen method, that is by using reading scales. Only "c" remains for which a measure has not been found. One has not been discovered by the writer in his reading, but quite possibly as Terence said of old, "There is nothing said which has not been said before." A method of measuring this knowledge is suggested in the following paragraph.

The first thing that a teacher should do in attempting to test achievement along this line is to make a survey of what books his students are able to obtain, for the extent of reading in science that a student may be expected to reach can be measured in all fairness only in terms of the books that he has at his command. For most, this will mean the school library, and the public library if there be one. Sometimes in the smaller schools the teacher might find it necessary to supplement the school library, as so many do now, with their own books. Other books may be obtained on loan from certain distant libraries. Obtaining books is really another problem outside the scope of this report. One practical way of measuring this growth is to prepare a list of books and magazine articles at the beginning of the year, post it up in the library or class room, or otherwise let the students know, then examine them at the end of the term.
or year on their knowledge of sources or general reading experience in
science based on the contents of these books.

The first type of question is a test of knowledge of sources taken
from the example in Hawkes, Linquist, and Mann, "The Construction and Use
of Achievement Tests, page 232;

DIRECTIONS:
In each of the exercises below, you are to suggest the sources which you
think best for getting information on the question given. Be as definite
in your suggestions as you can be. If you mention a book, or magazine, or
newspaper, state its name. If you do not know its title, tell how you
would find it. If you suggest some other sources, be just as definite in
describing them.

1. Where you could find out about the general principles which help to ex-
plain the methods of sending pictures by wire?

2. Where could you determine the relative electrical conductivity of iron
copper, and aluminum?

3. If you are making a special report on the corpuscular theory of light,
where would you get helpful information?

Such an exercise as this is not very objective and needs a very
definite set of directions to evaluate responses.

Purpose of test: (given in text, more or less as above.)

Values to be assigned to answers:
Allow 4 points credit for each source listed by the pupil which is
reliable for the information sought and which is as available as any
other equally reliable source.

Allow three points credit for each source listed by the pupil which
is reliable for the information sought but is not as available as
other reliable sources which he failed to mention.

Allow two points credit for each source which is likely to contain
some of the information sought but is only fairly dependable.

Allow one point credit for each source which is so vaguely defined
by the pupil that he would be unlikely to find it without considerable
loss of time. Also, one point for an omission or an "I don't know"
answer.

Allow no credit for any source listed by the pupil which is unlikely
to provide any helpful information or which provided information
which is not dependable.
A refinement of this type of question that suggests itself seems to be more objective and more easily marked. Although it demands the evaluating of sources, these sources are not likely to come from many points and make the marking a herculean job. If the student knows the sources of his reading material from the prepared lists this type of test item would seem to be quite fair as it measures achievement in a restricted field.

DIRECTIONS:
Below is a list of books and magazine articles with their authors and below each book is a suggestion or problem that may or may not be found in each book.
If you think that the book contains the suggestion or problem expressed in a manner that is moderately easy to understand and is reliable
Place a 3 in the blank provided.
If you think the book contains the suggestion or information but is very difficult to understand (too many technical terms and ideas)
Place a 2 in the blank provided.
If the source mentioned contains the information but the information is very unreliable or open to question
Place a "1" in the blank provided.
If the book does not contain that type of information at all
Place a "0" in the blank provided.

--------a. Beneath Tropic Seas (Wm. Beebe)
a very interesting book dealing with sunken treasure ships and the salvaging of these. Should be read by everyone.

--------b. The Hunger Fighters (Paul de Kruif)
tells among other things how Saunders of Ottawa developed the Marquis wheat for northern Canada.

--------c. The Universe Around Us (Sir Jas. Jeans)
explains in readily understandable English the marvels of the heavens.

--------d. The National Geographic Magazine
One issue contains a fine exposition on the building of an amateur's radio set.

--------e. 100,000,000 Guinea Pigs (Kallet and Schenck)
tells how the American public has hundreds of patent medicines foisted on it.

--------f. Creative Chemistry (Slosson)
a chemistry text that goes into the theoretical problems concerning the union of atoms in chemical reactions, with many mathematical equations to prove statements.
Marking responses:

As the questions stand now there is only one correct answer, which is evaluated in terms of the scale provided. The method of scoring is to give one point for correct responses and none for omissions and incorrect responses. However, along with the newer idea that is creeping into the evaluation of responses provision may be made for the pupil's recognition of the fact that he does not know and does not attempt to answer, in preference to giving him the same mark as one who makes a wrong answer. Three could be given for the correct responses, and one for blanks or "I do not know". In such a case this last response should be included in the scale of values presented to the student.

The matching type of question is quite useful here. An example follows:

DIRECTIONS: From this list of books select the letter in front of the name of each book that will associate the book with the topic or information in the list below:

a. Beneath Tropic Seas (Beebe)  g. Harvest of the Years (Burke, Wilbur)
b. Iron Paddlers (Davis)  h. Hunger Fighters (de Kruif)
c. Universe Around Us (Jeans)  i. Creative Chemistry (Slosson)
d. 100,000,000 Guinea Pigs (Kallet & Schenck)  j. Modern Physics (Dull)
e. Microbe Hunters (de Kruif)  k. How to Know Our Miner-
f. The Pet Book (Comstock)  l. Down Below the Light Zone (Williams)

1. Explains to you many of the fraudulent practices in preparing and selling many of the patent medicines in the United States of America.

2. Tells among other things of the general procedures of making popular perfumes, and their manufacture artificially.

3. Relates experiences in the first series of really deep-water exploration of sea-life.

4. Shows how the laws of physics apply throughout; explains man's greatest but most useless triumphs.

5. Explains how and why Marquis wheat was originated for northern Canada.
MARKING: This question type could be marked in the usual way in which matching questions have been marked, or evaluated on some scheme to credit "no response" as against actual wrong information or guessing. The correct response must always be valued at more than double the "no response" wherever any standard demands that the student reach an achievement of 50%, or some such arbitrarily set standard.

The multiple response type of question can also be utilized for the purposes of testing knowledge of sources.

1. You wish to find information how to make and sail a model sailboat. Check the sources mentioned below that you think most likely to contain this information and which you would go to first:

   ( ) Tours Through the World of Science
   ( ) The Scientific Study of Problems
   ( ) Installing Launch Engines
   ( ) Journal of Chemical Engineering
   ( ) Encyclopedia Britannica
   ( ) Popular Mechanics
   ( ) National Geographic Magazine
   ( ) Mechanic Monthly
   ( ) London Illustrated News
   ( ) L'Illustration

11. You wish to find a moderately recent article on the properties of matter under high pressure, yet not very technical. Write the name of the book or magazine listed above that you think most likely to contain this information.

These questions were tested on three classes. The tests seem to measure growth in reading experience, and indirectly growth in desire to read scientific literature. However no controls could be set up.

1. See Experiment V in the Appendix.
Benzene, which should not be confused with petroleum benzine, is formed during the fractional distillation of coal tar or similar products. Although it is an excellent solvent for gums, fats, and resins its most useful property is that it can be acted upon rather easily to form many other more important compounds, such as dyes, antiseptics, perfumes, and explosives. For many years after its discovery by Faraday, in 1825, its structural chemical formula remained unknown although its empirical formula had been established. This lack of knowledge materially hindered progress in the study of its chemistry and possibilities. In 1865 a German chemist who had long worried over this very problem sat down before his fire one evening after a hard day in his laboratory, and soon fell fast asleep. During his sleep his mind was occupied by a weird dream which conjured up snakes and atoms, writing around and in and out, and dashing madly about. "All at once", reported Kekulé the dreaming chemist, "I saw one of the snakes seize itself by its own tail, and the form then whirled mockingly before my eyes. As if by a flash of lightning I awoke and spent the rest of the night in working out the consequences of the hypothesis." His solution for the probable structural formula of benzene was the now famous benzene ring or hexagon composed of six carbon atoms, one at each angle, with an atom of hydrogen attached to the outside of each atom of carbon. This solution produced wonderful results in the study of benzene derivatives. Years after this discovery, Kekule, in commenting on the weirdness of it, declared, "We must learn to dream.

Check those statements thus ( ) which are either made directly in the paragraph or are reasonably deducible from the material. Do not check any one that is not deducible directly, nor made in the paragraph.

( ) 1. Dreaming is a very effective way to solve the great problems met in chemistry.
( ) 2. Benzine is prepared by distilling wood or coal tar.
( ) 3. The chief useful property of benzene is that it can be converted moderately easily into many derivatives.
( ) 4. Benzene is used mainly as a solvent for fats, resins, and gums.
( ) 5. Kekule discovered benzene.
( ) 6. The benzene ring as worked out by Kekule is:
SCORING: A not too cumbersome method of scoring would be to modify van Wagenen's method by totalling all the correctly placed check marks and the proper blanks left blank, then subtracting the check marks in the wrong places and the blanks which should have been checked. Before items of this type should be submitted for achievement testing they should be tested for difficulty and then graded. Those used on achievement tests should be of an approximately equal difficulty each time used in order that comparison of achievement will be constant.

Another variation for testing the ability to understand the literature of a subject is the use of multiple choice questions on a given paragraph as done this year for the first time on the Social Studies matriculation examination. An attempt will be made to adapt this technique to a physics example here:

"One is of course sufficiently familiar with the compression of a gas into a small volume when pressure is exerted on it, as dramatically demonstrated in converse by the air that rushed out of one's tire when one has a puncture, but the compression of liquids and solids is not so evident or so readily demonstrated. This has led to sometimes fantastic popular ideas about the absolute incompressibility of liquids like water, ideas which were supported by early crude experiments by physicists. Nevertheless, both solids and liquids, as well as gases, are compressible; the difference is merely one of degree, requiring much more delicate apparatus to disclose it. It is even more difficult to demonstrate the compressibility of solids; iron, for example, is 100 times less compressible than water, however, when pressure of thousands of atmospheres become available, the volume changes of liquids and solids become large enough to be measured accurately with comparatively simple means. Liquids may lose 30 to 40% of their volume. The volume of ice at room temperature at 50,000 atmospheres is found to be only 60% of the volume of the water with which the experiment started. Metals are in general much less compressible, but there is a great deal of variation, and the most compressible metal, is more compressible than ordinary liquids, and may be reduced to less than one-half its initial volume by a pressure of 50,000 atmospheres.

Two stages are to be recognized in the compression of a liquid and to a less extent in the compression of a solid. At first, while the pressure is comparatively low, the compressibility is comparatively high; this is followed at higher pressures by a relatively extended range of lower compressibility. The first stage is due to squeezing the atoms or the molecules into tighter contact-- "taking out the slack" from the atomic

1. van Wagenen, Reading Scales in Biology, Literature accompanying test (Manual).
structure. The second and more extended phase is due to more deep-seated changes which may affect the constitution of the atoms and the molecules themselves. The first stage can be understood with the stock of older ideas, which was adequate to explain the relations between liquids and gases, but to understand the more deep-seated alterations it is necessary to use some of the newer ideas of quantum theory.

Below is a series of questions based on this paragraph or excerpt. Check only those answers or statements that agree with the text given:

1. This report by Dr. Bridgeman of Harvard University states that
   ( ) a. gases are easily compressible at average conditions,
   ( ) b. gases can be compressed at only low temperatures,
   ( ) c. liquids are incompressible,
   ( ) d. solids can be compressed under tremendous pressures.

2. The compression of matter under certain conditions and enormous pressures
   ( ) a. permits at the lower levels the atoms and molecules to readjust their internal structure.
   ( ) b. can be completely understood on the basis of the familiar laws concerning the expansion and contraction of gases.
   ( ) c. causes it to pass through a series of events that can be grouped into two stages.
   ( ) d. forces the deep-seated alterations of molecules and atoms.

3. From this article a person could truthfully surmise that:
   ( ) a. the older scientists were so careful in their work that modern techniques really cannot improve upon theirs.
   ( ) b. scientists ideas of the laws of nature are not fixed but change with increasing experimentation.
   ( ) c. truth is not fixed but is variable.
   ( ) d. scientists concern themselves deeply with the most useless of things.
   ( ) e. further knowledge gained from the experiments might give us a better idea of the nature of the interior of our earth.

SCORING: This exercise could be scored simply by counting the correct scores. It would seem better to deduct mistakes, but perhaps not omissions. This is not an exact true-false situation where only one statement is made without any corroborating evidence. There is an element of multiple choice in the multiple response. Both the practice of deducting errors and of not deducting them are followed on standardized tests for this type of question.
CONCLUSIONS AND RECOMMENDATIONS.

1. It is axiomatic that testing and teaching should bear directly upon the objectives of the course. The objectives must be kept in mind quite clearly when tests are being prepared.

2. As a corollary to the first proposition it would seem that the objectives of a course must be selected with great care. The objectives of a course should not be restricted to only those wherein achievements can be measured readily. If an objective is attainable, is deemed good, and is worth the effort and time, then it must be included, and testing programmes must be adapted to the measuring of this objective.

3. It seems highly desirable that the credits or marks recorded to the student’s account should be apportioned approximately in the magnitude of the percentage weightings found by means of the questionnaire. (See Chapter 11).

In referring to these objectives and their evaluations it will be seen that the first one pertaining to factual knowledge is given a weighting of twenty percent. Compared with the weighting of sixty-five to ninety percent that many science courses give to information the twenty percent seems very low. However, to this twenty percent, which should be construed as mainly "pure" scientific information may be added the ten percent of questions referring to health. Another seven percent directed to testing the student’s knowledge of the contributions of science to civilization can be combined with the others, together with seven percent to be used in counteracting false beliefs and superstitions.
All these total forty-four per cent, or approximately fifty percent. That is, fifty percent of the student's credit should be allocated to these objectives. This does not appear to be quite so drastic a reduction as the bare twenty percent appeared because many achievement tests do include some or much of these additional types of items.

May it be said in passing that it is this type of material subtended under these objectives which is most easily "crammed" by students for writing tests and examinations.

The adoption of a procedure similar to that advocated here would tend considerably to discourage "cramming", or at least to devaluate it.

4. It is highly desirable to break the testing programme into four groups of tests. Similar suitable testing procedures apply to certain objectives. Other groups of objectives present similar difficulties in testing. It is thought advisable to segregate the objectives into the following groups in order to simplify the various complications:

(a) Informational; (b) Vocational and Avocational; (c) Reading, and (d) Performance (or "Practical") Tests and Scientific Method.

(a) Informational

Regular unit or achievement tests can be made to include the four objectives I, IV, VI, VII (body of scientific knowledge, Health, achievements in the field of science, superstitions and false beliefs) in the proportion of 20, 10, 7, 7. To round the weighting slightly in order to give a total of fifty the parts would then read, 25, 10, 8 and 7. These tests can be administered as usual following the completion of units of work, or on a final achievement test.
(b) Vocational and Avocational

Tests on the worthy use of leisure (objective V) and exploring vocations (objective VIII) could be combined in the proportion of 6 to 10 respectively. It might be argued that these could be placed on achievement tests because pencil and paper tests can be prepared for these objectives. To obtain a reliable measure requires questions of a sufficient number that the testing will occupy a period of class time. For this reason it was thought advisable to separate this material from the preceding. These tests can be given at the end of each unit. It would seem better to give them four times a year instead, because more comparative work between units could be included. The testing would not be too restricted then.

(c) Reading

Reading tests and measuring increased desire to read (Objective X) can be fitted in to the ordinary units. It is possibly desirable to include one of these reading ability tests in each unit of work. They can be taken from test books available and be directly on the work so that no time is lost, e.g., silent reading tests. The "desire to read" should be checked at the beginning and end of a year's work.

(d) Performance and Scientific Method

The "performance" type tests would include some under objective II (scientific method), III (resourcefulness) and IX (experimentation) in the proportion of 18, 11, 7. Testing achievements under these objectives could not be done very frequently. Twice per year is likely to be the best number of times to test. If these tests are to be administered twice per year the first should occur certainly no later than January and should
be looked upon largely as a diagnostic test to see wherein the pupils are weak. The final test should be given after all the units are completed because each unit of work will likely add some additional procedure to their knowledge of the scientific method, or at least some new sidelight. The contributions of each unit to resourcefulness should be important, and the ability to experiment should improve with each unit of work taken.

It would seem advisable to do this testing at the end of a course in order to obtain better comparative test items utilizing procedures in several tests. For schools that make a practice of promotion by unit there would be many new problems. If such schools tested all objectives at the completion of each unit with any degree of reliability it would take nearly a five-period week to do it. This would seem to be a very great amount of time to spend on testing. If they did not test all achievements they would be falling down in their duties. If they left these general achievements to the end of a series of units (or a course) they might find that they are in a dilemma, for it is conceivable that a student may reach a satisfactory level of achievement on the informational test but be quite unsatisfactory in his ability to use the scientific method and to experiment. What should be done with such a student? He has "passed" in half his work and "failed" in the other (which is not taught as a distinct unit for him to repeat).

These three objectives are included together because the type of testing will, or should, be much alike for all. There should be actual manipulation in many cases, which would necessitate special arrangements in the laboratory or some other room. These arrangements take considerable time and cannot be made with the facility that pencil and paper tests can be.
The sampling in these tests must be as extensive as time will permit in order to gain as much reliability as possible.

5. For a simplified introductory programme a ratio between the "informational" and "practical or doing" objectives of forty-four to thirty-six (five to four) could be maintained.

For the first year if the examiner only attempted to follow out these two major means of measuring he will have made a drastic break, and a very valuable one, with traditional methods of measuring achievements.

It possibly is advisable that the examiner keep to these two fields for his first attempts, or at least until he feels he is on safe ground.

6. Tests involving diagrams, graphs, scales, etc., should be used in testing the factual side of learning more than they are used at present on tests. They can be made very useful.

7. Diagrams should not be considered as a substitute for any test demanding activity or manipulation of materials. There are common elements between these two methods of testing but they vary too much between individual techniques, and correlations are too low even for the highest to warrant the assumption of equality. Because of the usefulness of diagrams on achievement tests it should not be assumed that they are valid tests for resourcefulness, experimentation and similar abilities. (For further information see Appendix pp. XIV to XXVI.)

8. Increase in reading interest of students in Science can be measured from grade to grade. Science teaching does produce increments of interest among pupils in extra-curricula reading. (See Experiment V in Appendix p. XXVII). Again, as elsewhere, it is the direct attack upon the stimu-
lation of interest that produces best results. By direct attack it is not meant that all books be assigned, but that reading should be as free as possible. It may be necessary to assign a few books to start off the programme, but this assignment should be reduced to a minimum.

9. It seems within the realm of possibility that tests can be made to measure all achievements or abilities. These tests and questions have been suggested in chapters V, VI, and VII.

10. Interest of students in avocational science grows with each course taken. This growth can be measured (See Appendix, Experiment I, p 1). The survey of standardized tests reported in Chapter IV would seem to suggest that no present standardized tests would be satisfactory for testing in General Science IV or V, nor would a battery of present tests suffice.

11. The form of the question should be such that it tests the mental process desired: memory (recall, completion in the main), association (matching), selection (multiple choice, multiple response), deduction (from data supplied), computation, comparison. Factual or informational tests will use the first three forms mainly. Testing the ability to use the scientific method would make use of the last four to a greater extent. Testing should be done as much as possible in science for the understanding of the basic principle by emphasizing the "why", "the consequences of such an action", "how does it happen" ideas.

12. The validation of the testing of the ability to use the scientific method will be rather difficult. Results of tests will need to be compared with the teacher's opinion of the student's ability, with his
utilization of the method in his daily life after he leaves school, with the success he has in using the method in university work if he goes there. The most valuable criterion might be the extent to which the scientific method is applied in solving the problems of daily life. The objective really aims for this. To measure this transfer from the teaching in school to the application in daily life is a task too extensive for one teacher to do. It really needs to be undertaken by some central authority.

Likewise the validation of the tests on vocational and avocational aspects of science courses can be obtained only by seeing to what extent the results of the tests have guided students into or away from scientific vocations and avocations. This validation would demand a considerable "follow up" programme.

13. The purpose of the test must be kept clearly in mind. The teacher must ask himself, "Is this to be a mastery test whose chief purpose is to discover if basic principles have been grasped, an achievement test to see how much of details and basic principles has been learned, or an administrative test to rank or place the student or to secure data?"

14. There should be a clearer, perhaps a new, connotation given to mastery tests, achievement tests, and administration tests. To study the function of each as it should be is really another problem.

It is suggested that mastery tests be directed toward the measuring of basic achievements that no one is likely to forget. The extent of this area cannot be stated until measurements are taken for each course. Region A on the chart represents this area. (Figure X:111 next page)

Administrative tests are composed of only those items which have
good discriminative power. These items, the more difficult, are represented by region B on the chart.

Achievement tests should be made to measure both fields, the discriminative items and the basic mastered items of regions B and A.

![Achievement Tests Diagram](image)

The achievement test really should be one to measure the pupil's growth. As it is used and interpreted rather widely at the present time it is really an administrative or grading test from which A's, B's, C's, D's, E's are obtained.

15. The present connotation given to achievement tests implies the evaluating of all achievements on the basis of a single composite score for each test.

A grave philosophical and mathematical problem arises here which cannot be answered in this work. Can these diverse abilities be measured and summated in terms of a single score? The theory of additivity demands that only addends of identical quality be summated. In other words it involves the old problem of trying to add five rocks and four oranges.

This difficulty is rather a serious one and needs to be faced soon. In connection with the reform of report cards of student progress the philosophical basis implied here must be settled. If a report card is
going to contain only a reference to total achievement in a subject there must be the summation of scores of possibly too diverse quality. If these achievements cannot be summated it would demand a report card with progress evaluated in terms of each objective for each course. If report cards are to be modified further until they become diagnostic reports it would be more important still that all testing be done in accordance with the objectives of the science course.
BIBLIOGRAPHY

Organizing the Problem.

1. Alexander, Carter; (Librarian Professor, Teachers College, Columbia University), How to Locate Educational Information and Data; Bureau of Publications, Teachers College, Columbia University; 1935.


Testing Procedures and Theory of Measurement.


5. Draper, E.M.; Principles and Technique of Curriculum Making; D. Appleton-Century; 1936; Chapter on testing, Chapter XV.


15. Pintner, Rudolph; *Intelligence Testing, Methods and Results*; Henry Holt and Co., New York; 1923.

16. Programme of Studies for the High Schools of British Columbia; Bulletin 1, pp. 5-26, 159-189.


BIBLIOGRAPHY.


PHYSICS


CHEMISTRY


3. Brownlee, R.B., Fuller, Hancock, Schoon, Whitst; First Principles of Chemistry; Allyn and Bacon; 1934.

4. Bruce, G.H.; High School Chemistry; Allyn and Bacon; 1931.


8. Jaffe, Bernard; New World of Chemistry; Silver, Burdett, and Co.; 1933.
BIBLIOGRAPHY


BIOLOGY


3. Cole, E. C.; An Introduction to Biology; John Wiley and Son; 1933.


BIBLIOGRAPHY

BIOLOGY.

17. Wood, G.C., and Carpenter, H.A.; Our Environment; Allyn and Bacon; 1934.

GENERAL SCIENCE.

1. Allen, T.B., and Ferguson, W.F.; Elementary Science; Ryerson Press; 1934.
WORKBOOKS

1. Adell, J.C., Dunham, O.O., and Walton, L.E.; Explorations in Biological Science; Ginn and Co.; 1937.


5. Brownlee, Fuller, Hancock, Sohon, Whitsit; Laboratory Experiments in Chemistry; Allyn and Bacon, 1935.


11. Fuller, Brownlee, and Baker; Laboratory Exercises in Physics; Allyn and Bacon; 1932.


BIBLIOGRAPHY

TESTS, STANDARDIZED AND OTHERWISE.

BIOLOGY

1. Cooprider, J.L.; Cooprider Biological Information Test; 1924; Public School Publishing Co., Bloomington, Ill.

2. Downing, E.R., and McAtee, Veva M.; Biology Unit Tests; Arrangement A and B; 14 tests; Lyons and Carnahan.


4. Ruch, G.M., and Cossman, L.S., Ruch-Cossman Biology Test; Forms A and B; 1924; World Book Company, 2126 Prairie Ave., Chicago.

5. Van Wagenen, M.J.; Van Wagenen Reading Scales in Biology; Forms A and B; 1929; The Educational Test Bureau, Inc., Minneapolis, Minn.

CHEMISTRY


3. Gerry, H.L.; Harvard High School Chemistry Test; Forms A and B; Givin and Co.; 1922.


5. Malin, J.E., Diagnostic Test in Chemistry; Forms A and B; Public School Publishing Co., Bloomington, Ill.; 1932.

6. Persing; Laboratory Chemistry Test; Forms A and B.

7. Powers General Chemistry Test; Forms A and B; World Book Co., 2126 Prairie Ave., Chicago.
BIBLIOGRAPHY.

PHYSICS

1. Cooperative Physics Test; Cooperative Test Service; New York.

2. Farwell, H.W., Wood, B.D.; Columbia Research Bureau Physics Test; Forms A and B; 1926; World Book Co.

3. Hughes; Physics Scales, Information and Thought; Public School Publishing Co., Bloomington, Ill.


GENERAL SCIENCE.


2. Cooperative General Science Tests (Underhill and Powers); Forms 1936, 1937; N; Cooperative Test Service, New York.

3. Dvorak, A.; General Science Test; Forms S₂, T₂, R₁; 1924; Public School Publishing Co., Bloomington, Ill.


5. Powers, S.R.; General Science Test; Forms A and B; 1927; Bureau of Public Teachers College, Columbia University.

6. Ruch and Popenee; General Science Test; Forms A and B.

OTHER TESTS

1. Brewer; Vocational Aptitude Test for Boys;


3. Detroit Mechanical Aptitude Test.


5. Stewart-Ashbaugh Physics Tests, Form 1; Public School Publishing Co., Bloomington, Ill.


7. Sangren and Marborger; Michigan Institute Test in Physics;
APPENDIX

EVALUATING THE OBJECTIVES OF THE GENERAL SCIENCE COURSES IV AND V.

PURPOSES OF QUESTIONNAIRE.

It is desired to find an evaluation of the objectives of the present courses in General Science IV and V in order to guide the development of question types that would be suitable for achievement test and valid for the objectives of the courses. Testing would also involve some means of apportioning achievements towards the various objectives, that is, a weighting must be established.

EVIDENCE OF NEED OF QUESTIONNAIRE.

1. Programme of Studies for High Schools of B. C., 1937.
   Page 27 re matriculation and accrediting: "A system of accrediting will be established. There is, therefore, no longer any reason why high school teachers in their teaching and school procedures should hesitate to aim at the achievement of the general objectives of education and particularly at those objectives which may be achieved through their own subjects. This applies equally to methods of testing, for it is fundamental that testing should bear upon the objectives of a course.
   Page 21. An examination that stresses these outcomes (knowledge and skill) to the exclusion of the others not only excludes these outcomes from the testing procedures, but speedily excludes them from the teaching procedures as well.

   "The moment we attempt to apply a systematic procedure to our teaching we are confronted with the need of a process of testing; It is not enough to teach; it is necessary to find out whether our teaching has registered.

   (paraphrased)
   The results of this experiment apparently support the claim that direct methods of attack result in most accomplishment. It would seem that the best way to secure a desired result is to teach for that result

   This experiment leads, a priori, to the conclusion that there must be definite aims in the course in Physics, or any other course.

4. Thorndike, E.L. His theory that transfer of training occurs only insofar as there are identical elements in both fields that are consciously realized would suggest that any course must be taught toward the objectives of that course. Therefore again, evidence from his experiments and investigations would commit a teacher to teach directly for anything which is to be done successfully.
   Some sort of test would be the only way of finding success or failure.

5. Hawkes, Linguist, and Mann. "The Construction and Use of Achievement Tests" Page 5. One major defect of typical examinations has been the fact that they have given evidence with reference to only a limited
number of objectives of that course, and have not indicated adequately the degree to which students were attaining all the desired outcomes of instruction.

Page 7. Rarely do we find students tested on such objectives as their ability to utilize the scientific method, consistency of point of view, their skill in laboratory work.

Page 13. It is readily apparent that the procedures of formulating and analyzing the major objectives for any course are desirable, and are invaluable when making a comprehensive programme of examinations.

DIRECTIONS FOR EVALUATING THE OBJECTIVES.

Certain objectives have been laid down by the Committee which prepared the courses General Science IV and V. It is desirable to find which objectives various groups of people think most important. Two means of evaluating these would be:

1. to rank them numerically in order of importance 1 to 10 as you see fit

and

2. to give a percentage rating on the basis of an estimate as to what fraction of the total outcomes each objective is worth.

Below are found the ten objectives selected by the Committee, but jumbled in arrangement lest there was any attempt by the Committee at weighting. (In order to avoid outside influences upon your decision.) To the right of each statement there are two columns. In Column "A" place your numerical or ordinal ranking values giving the value 1 to that objective that you think most valuable in the course, 2 to the next, and so forth up to 10. In Column "B" place an estimated percentage of "worth" beside each objective. (It is admitted that up to the present no very satisfactory tests have been developed for some of these objectives, so please evaluate these items on the idealistic assumption that valid tests can be prepared for each. Such tests might be of a quite different kind from the usual pencil and paper type.)

Each responder is asked not to refer to books or to other person's opinions, but to use his own careful judgment as to what he thinks best.

On the questionnaire although a space is provided for your name it is not necessary to fill this in if you do not wish. However, an identification would add interest.

(questionnaire on next page.)
**QUESTIONNAIRE (Submitted to teachers)**

**NAME**

**HAVE YOU TAUGHT SCIENCE SUBJECTS?**

If so which of these? Agriculture, Biology, Chemistry, Gen. Science, Physics

**DO YOU CONSIDER THE FIELD OF SCIENCE TO BE YOUR SPECIALTY?**

<table>
<thead>
<tr>
<th>OBJECTIVE</th>
<th>RANK PERCENTAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. To provide materials for the worthy use of leisure.</td>
<td></td>
</tr>
<tr>
<td>b. To enable the student to counteract superstition and to correct erroneous beliefs through the application of scientific principles.</td>
<td></td>
</tr>
<tr>
<td>c. To develop resourcefulness and adaptability to new conditions.</td>
<td></td>
</tr>
<tr>
<td>d. To acquire a body of knowledge in the fields of science which will enable the student to interpret and appreciate his environment.</td>
<td></td>
</tr>
<tr>
<td>e. To appreciate achievements in the field of science and the contributions of scientists to the world.</td>
<td></td>
</tr>
<tr>
<td>f. To develop ability in the use of the scientific method, e.g.,</td>
<td></td>
</tr>
<tr>
<td>a. To make accurate observations and to record them systematically.</td>
<td></td>
</tr>
<tr>
<td>b. To draw valid conclusions.</td>
<td></td>
</tr>
<tr>
<td>c. To suspend judgment until sufficient evidence has been obtained.</td>
<td></td>
</tr>
<tr>
<td>d. To develop a critical yet tolerant attitude towards new ideas.</td>
<td></td>
</tr>
<tr>
<td>g. To develop the desire to read scientific literature</td>
<td></td>
</tr>
<tr>
<td>h. To acquire knowledge that will contribute to public and personal health.</td>
<td></td>
</tr>
<tr>
<td>i. To explore the field of science in order to assist the pupil to choose his vocation.</td>
<td></td>
</tr>
<tr>
<td>j. To develop the ability to perform simple experiments and thus to appreciate the scientific basis of science.</td>
<td></td>
</tr>
</tbody>
</table>
EXPERIMENT I

Purpose. To try to develop a testing procedure which will give a measure of the avocational interest in science created by a particular course in science.

Preparation of Tests. Lists of possible avocational activities were prepared by selecting those suggested in a wide variety of text books. The lists were separated roughly into sections based mainly physics, chemistry, or biology. Each item was to be evaluated by the student on a seven column scale as shown in the sample. To reduce the mechanics of student effort, the student was asked to make only a check mark in the proper column.

Administration. The test was given to one class from each of Grade IX, X, and XI in Britannia High School. These classes were 17, 10, and a mixed group taking General Science V. It was not possible to alter the classes at all in order to equate them for ability (either in general or in science) or interest. The Grades X and XI classes are approximately average as shown by the school records of the students. The Grade IX class (Class 17) may be slightly above average, because their age grade average is lower than the normal of 14 years taken on September 1, 1938. So far no rankings or tests of the class have been made.

Each class was given sufficient time to finish the test. The slowest students usually required thirty-five minutes.

Scoring was done simply by adding the check mark in each column. This was extremely easy and is one of the advantages of this particular form of test. Scores for each column were then averaged.

<table>
<thead>
<tr>
<th>TABLE SHOWING RELATIONSHIP BETWEEN INTERESTS DEVELOPED AND NUMBER OF SCIENCE COURSES TAKEN.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last Year's Science Course</td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>GRADE IX</td>
</tr>
<tr>
<td>Class 17</td>
</tr>
<tr>
<td>GRADE X</td>
</tr>
<tr>
<td>Class 11</td>
</tr>
<tr>
<td>GRADE XI</td>
</tr>
</tbody>
</table>
CONCLUSIONS

Several warnings must be made, however, before conclusions are drawn because:

1. The data are based on only three classes of different ages.
2. No control group which had no science training could be found.
3. Classes were not equated although they were roughly "average".

The results seem to show that the test has the power to distinguish increments in interests and avocations attributable to science courses taken by the students.

Each year more students see the possibility of more science avocations. The increase in possible avocational activities amount to four or five items. Actual avocational activities increase each year being more pronounced in Grade IX General Science III. The reason for this is not shown on the test, but it may be due to the fact that students begin to handle things more and to perform experiments more by themselves in this and later courses (compared to conditions in elementary schools).

Avocational interests maintained increase between General Science III and IV, but dropped off for General Science V. This may be due to the greater demands of academic work and courses of the senior years of the high school courses, made upon the student. Other factors may be operative.

There is an increase in the number of avocational activities dropped each year. This may be a natural concomitant of mental growth for few persons maintain every avocational activity that they adopt.

With increases shown in interests and hobbies each year, it is only to be expected that fewer neutral or "no interest" responses should be made each year.

It was rather surprising to the investigator that so very few activities were dropped due to the negative influence of the science courses taken, as a slightly higher average was expected.

Few students think that any of the activities of the list are harmful or should not be adopted.

GENERAL

The same test when given to individuals in later years might reveal an individual growth of interests comparable to the group increments mentioned. This could not be tested by the investigator at the present.
AN EXERCISE TO SEE HOW SCIENCE HELP YOU TO USE YOUR LEISURE TIME.

As the title suggests this exercise is set as an attempt to find out to what extent science courses have provided you with worth-while interests activities, and hobbies to follow in your leisure time.

In reading through this exercise you will see first a list of values or marks to give certain items in the second list which follows it. In order to save your time as much as possible you are asked to put check marks only in the columns that fit your evaluation best. After you have read each item carefully place a check mark in

Column 5 to represent activities, interests which you have not followed but would like to some time if you have sufficient funds and leisure

Column 4 to represent those hobbies, activities, interests, etc., that you have adopted or have developed during this course.

Column 3 to represent those hobbies, activities, interests, etc. that you have been following for some time, and still continue to do so.

Column 2 to represent those hobbies, activities, interests, etc. that you did follow at one time but dropped before you took this course.

Column 1 to represent those activities in which you think a student should not participate at all.

Column 0 to represent those activities in which you have no interest, and still have no interest.

Column -1 to represent hobbies, interests, activities, etc. that you followed at one time but have dropped or turned against as direct result of this course. Please consider these fairly.

EXAMPLE:

<table>
<thead>
<tr>
<th>ITEM</th>
<th>GROUP 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Growing prize-winning flowers</td>
<td>5 4 3 2 1 0 -1</td>
</tr>
<tr>
<td>2. Studying the stars</td>
<td></td>
</tr>
<tr>
<td>3. Collecting diamonds by stealth</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ITEM</th>
<th>PLACE EDGE OF CHECK SHEET HERE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Devising lighting effects</td>
<td>1.</td>
</tr>
<tr>
<td>3. Constructing small telephone systems other electrical things</td>
<td>3.</td>
</tr>
<tr>
<td>4. Making small electrical motors for boats, etc.</td>
<td>4.</td>
</tr>
<tr>
<td>5. Experimenting with static electricity</td>
<td>5.</td>
</tr>
<tr>
<td>6. Using the photo-electric eye in various hook-ups</td>
<td>6.</td>
</tr>
<tr>
<td>8. Making radio sets of simple or complex types</td>
<td>8.</td>
</tr>
<tr>
<td>10. Mirrors and &quot;magic&quot; with them</td>
<td>10.</td>
</tr>
<tr>
<td>11. Experimenting with prisms and lenses</td>
<td>11.</td>
</tr>
<tr>
<td>12. Photography; color and lighting effects</td>
<td>12.</td>
</tr>
<tr>
<td>16. Renovating old cars, doing all work yourself</td>
<td>16.</td>
</tr>
<tr>
<td>17. Constructing model steam engines</td>
<td>17.</td>
</tr>
<tr>
<td>18. Doing tricks, etc. based on the knowledge of inertia, centre of gravity, equilibrium</td>
<td>18.</td>
</tr>
</tbody>
</table>
19. Experimenting with heat, expansion and contraction...19.
20. Devising machines to harness various forms of energy...20.
22. Projection lanterns..............................................22.
23. Making model airplanes..........................................23.
27. Sailing and building model yachts................................27.
28. Studying the stars, directions, articles on astronomy.........28.
29. Reading books and articles on electricity etc...................29.
30. Reading books and articles on radio and wireless..............30.
31. Reading to keep up with modern discoveries in physics....31.
If you have other interests, etc. add them here and treat likewise 32.
33.
34.
35.
36.
37.

GROUP II

1. Glass bending and blowing........................................1.
3. Releasing gases from compounds..................................3.
5. Studying "First Aid" for poisons of all kinds...................5.
7. Collecting minerals and ores.....................................7.
8. Trying to smelt and refine ores and minerals...................8.
11. Testing the action of electricity on compounds and elements...11.
12. Making and collecting crystals of different kinds........12.
14. Tracing the chemical properties of families of elements...14.
15. Making your own paper for blue-prints, exposing, developing...15.
16. Experimenting with sulphur......................................16.
17. Making modern resin such as bakelite, redmanol.............17.
19. Performing "tricks of chemical magic"..........................19.
20. Experimenting with the spectacular elements..................20.
22. Visiting Chemical industrial plants................................22.
27. Making explosives..................................................27.
28. Fuels, sampling and experimenting with them..................28.
29. Preparing diagrams of industrial chemical processes......29.
32. Collecting and studying textiles................................32.
33. Studying anaesthetics........................................33.
34. Perfume making...............................................34.
35. Distillation..................................................35.
36. Analyzing patent medicines, fake "drugs"...............36.
37. Making soaps................................................37.
38. Making your own sprays for spraying garden plants
    and trees.................................................38.
40. Reading books and articles on modern chemistry.......40.

If you have other special interests of a chemical nature that are
not included on the list above add them below and treat similarly.
41.
42.
43.
44.
45.
46.
47.
48.
49.
50.
1. Developing hybrid plants
2. Breeding experiments on Mendel's Laws
3. Experiments on plant nutrition
4. Grafting, budding and layering
5. Going on nature hikes
6. Growing plants under various conditions to test reactions
7. Trying to train low forms of animal life
8. Making and maintaining a balanced aquarium
9. Attracting desirable birds around your place
10. Building and caring for an ant's nest, or a demonstration nest
11. Maintaining a demonstration hive of bees
12. Growing protozoa and bacteria in a hay infusion
13. Carrying on regeneration experiments with hydra crabs, etc.
14. Controlling some insect pest of your region
15. Photographing "nature in the Wild"
16. Collecting harmful parasites, preserving
17. Collecting useful parasites
18. Collecting weeds
19. Collecting seasonal flowers (spring, summer, fall)
20. Collecting families of flowers (rose, lily, composites, heather)
21. leaves of trees native to region
22. leaves of introduced trees
23. bark and wood samples of various trees
24. ferns of district
25. mosses
26. mushroom and toadstools, etc., spore prints
27. fungal parasites, etc.; moulds, wildews
28. dry fruits
29. juicy fruits
30. how plants defend themselves
31. fossils
32. plants to show means of seed and fruit dispersal
33. insects to show orders, families, etc.
34. insects to snowlife histories; Riker mounts
35. Aquatic or marine forms of life
36. Insect galls on plants
37. examples of protective coloration
38. Making life history studies of insects (moths, butterflies, beetles, etc.)
39. studies of some insect pest or plant parasite
40. studies of chicken or some bird
41. studies of frogs or toads
42. studies of mammals
43. Observing and identifying all the birds of your locality.

44. Observing and identifying all the mammals of your locality.

45. Studying the various plant and animal habitats and environments.

46. Studying alpine or mountain associations.

47. Mapping all the trees in a given area, etc.

48. Mounting birds or mammals (taxidermy).

49. Reading stories about wild animals, accounts.

50. Reading general biological topics.

If there are interests of yours that have not been mentioned here add them to the list in the spaces below and treat them similarly.

51.
52.
53.
54.
55.
56.
57.
58.
59.
60.
## INTERESTS OF SCIENCE STUDENTS

### SCIENCE III  CLASS 17

<table>
<thead>
<tr>
<th>COLUMNS</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, J</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>103</td>
<td>0</td>
</tr>
<tr>
<td>C, M</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>107</td>
<td>0</td>
</tr>
<tr>
<td>C, LeR</td>
<td>35</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>71</td>
<td>1</td>
</tr>
<tr>
<td>D, D</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>L, I</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>113</td>
<td>0</td>
</tr>
<tr>
<td>L, F</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>111</td>
<td>0</td>
</tr>
<tr>
<td>L, B</td>
<td>13</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>L, S(D)</td>
<td>14</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98</td>
<td>0</td>
</tr>
<tr>
<td>M, H</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>103</td>
<td>0</td>
</tr>
<tr>
<td>M, R</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>115</td>
<td>0</td>
</tr>
<tr>
<td>M, S</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>107</td>
<td>0</td>
</tr>
<tr>
<td>McA, F</td>
<td>27</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>9</td>
<td>77</td>
<td>0</td>
</tr>
<tr>
<td>McD, S</td>
<td>18</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>McG, J</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>121</td>
<td>0</td>
</tr>
<tr>
<td>McI, H</td>
<td>22</td>
<td>1</td>
<td>11</td>
<td>1</td>
<td>0</td>
<td>85</td>
<td>0</td>
</tr>
<tr>
<td>McK, JD</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>McW, A</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>107</td>
<td>0</td>
</tr>
<tr>
<td>N, M</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>N, E</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>P, T</td>
<td>32</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>84</td>
<td>0</td>
</tr>
<tr>
<td>R, M</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>104</td>
<td>0</td>
</tr>
<tr>
<td>S, M</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>0</td>
</tr>
<tr>
<td>S, B</td>
<td>16</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>109</td>
<td>0</td>
</tr>
<tr>
<td>S, C</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>S, D</td>
<td>15</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>0</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>S, A</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>109</td>
<td>0</td>
</tr>
<tr>
<td>T, H</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>T, Y</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>T, M</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>U, K</td>
<td>15</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>W, T</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>W, G</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>121</td>
<td>0</td>
</tr>
<tr>
<td>W, D</td>
<td>34</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>W, J</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>109</td>
<td>0</td>
</tr>
<tr>
<td>W, E</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>116</td>
<td>0</td>
</tr>
<tr>
<td>Y, T</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>99</td>
<td>0</td>
</tr>
</tbody>
</table>

546  14  67  50  23  3774  1
### INTERESTS OF SCIENCE STUDENTS

#### SCIENCE IV CLASS 10

<table>
<thead>
<tr>
<th>STUDENT</th>
<th>COLUMNS 5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A</td>
<td>13</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>B, H</td>
<td>12</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>B, A</td>
<td>27</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>0</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>B, D</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>106</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>C, J</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td>C, M</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>104</td>
<td>0</td>
</tr>
<tr>
<td>C, R</td>
<td>19</td>
<td>1</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>91</td>
<td>1</td>
</tr>
<tr>
<td>C, R</td>
<td>19</td>
<td>1</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>91</td>
<td>1</td>
</tr>
<tr>
<td>D, R</td>
<td>12</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>E, D</td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>G, A</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>98</td>
<td>0</td>
</tr>
<tr>
<td>H, S</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>98</td>
<td>0</td>
</tr>
<tr>
<td>J, C</td>
<td>17</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>K, T</td>
<td>16</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>K, W</td>
<td>44</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>70</td>
<td>0</td>
</tr>
<tr>
<td>L, V</td>
<td>14</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>105</td>
<td>0</td>
</tr>
<tr>
<td>M, N</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>87</td>
<td>0</td>
</tr>
<tr>
<td>N, H</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>87</td>
<td>0</td>
</tr>
<tr>
<td>P, E</td>
<td>26</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>P, W</td>
<td>10</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>P, M</td>
<td>10</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>R, T</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>R, J</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>R, H</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>109</td>
<td>0</td>
</tr>
<tr>
<td>R, M</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>0</td>
</tr>
<tr>
<td>R, N</td>
<td>26</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>85</td>
<td>0</td>
</tr>
<tr>
<td>S, J</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>99</td>
<td>0</td>
</tr>
<tr>
<td>S, J</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>99</td>
<td>0</td>
</tr>
<tr>
<td>S, E</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>S, E</td>
<td>11</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>S, G</td>
<td>13</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>T, M</td>
<td>16</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>105</td>
<td>0</td>
</tr>
<tr>
<td>T, P</td>
<td>57</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>52</td>
<td>0</td>
</tr>
<tr>
<td>W, F</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>L, R</td>
<td>77</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>108</td>
<td>0</td>
</tr>
<tr>
<td>D, M</td>
<td>28</td>
<td>1</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>80</td>
<td>0</td>
</tr>
</tbody>
</table>

<p>| 624 | 52 | 120 | 110 | 10 | 3580 | 3 |</p>
<table>
<thead>
<tr>
<th>STUDENT</th>
<th>COLUMNS 5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
<th>-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, T</td>
<td>13</td>
<td>30</td>
<td>14</td>
<td>14</td>
<td>0</td>
<td>52</td>
<td>0</td>
</tr>
<tr>
<td>A, E, J</td>
<td>21</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>97</td>
<td>1</td>
</tr>
<tr>
<td>A, N</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>B, J</td>
<td>29</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>76</td>
<td>0</td>
</tr>
<tr>
<td>B, D</td>
<td>37</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>C, Y</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>C, F</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>115</td>
<td>0</td>
</tr>
<tr>
<td>F, H</td>
<td>32</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>G, F</td>
<td>14</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>1</td>
<td>93</td>
<td>0</td>
</tr>
<tr>
<td>G, L</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>115</td>
<td>0</td>
</tr>
<tr>
<td>G, L</td>
<td>51</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>0</td>
</tr>
<tr>
<td>G, A</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>109</td>
<td>0</td>
</tr>
<tr>
<td>G, F</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>113</td>
<td>0</td>
</tr>
<tr>
<td>H, W</td>
<td>20</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>H, H</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>H, R</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>H, W</td>
<td>20</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>93</td>
<td>0</td>
</tr>
<tr>
<td>K, J</td>
<td>5</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>L, M</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>L, V</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>L, D</td>
<td>19</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>M, N</td>
<td>15</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>0</td>
</tr>
<tr>
<td>M, C</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>88</td>
<td>0</td>
</tr>
<tr>
<td>M, I</td>
<td>19</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>M, Y</td>
<td>13</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>M, O, C, B</td>
<td>41</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>78</td>
<td>0</td>
</tr>
<tr>
<td>P, M</td>
<td>57</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>S, F</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>107</td>
<td>0</td>
</tr>
<tr>
<td>S, M</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>108</td>
<td>0</td>
</tr>
<tr>
<td>S, A</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>S, A</td>
<td>32</td>
<td>3</td>
<td>31</td>
<td>0</td>
<td>51</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>T, S</td>
<td>49</td>
<td>5</td>
<td>10</td>
<td>2</td>
<td>53</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>W, S</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>108</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>W, H</td>
<td>31</td>
<td>1</td>
<td>16</td>
<td>3</td>
<td>70</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>W, T</td>
<td>66</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>W, J</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>W, E</td>
<td>31</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>84</td>
<td>0</td>
</tr>
<tr>
<td>S, M</td>
<td>25</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>Z, E</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>98</td>
<td>0</td>
</tr>
</tbody>
</table>

|        | 813 | 57 | 105 | 137 | 13 | 3550 | 2  |
EXPERIMENT II

PURPOSE To see how students' abilities to identify structures shown in diagrams of organisms correlates with abilities to identify the actual organs in the organisms.

PROCEDURE Three common and relative simple subjects were chosen, the clam, a gladiolus flower, and a cross section of a angiosperm tree trunk.

The diagrams, to match were prepared anew directly from an average sample of each type of specimen. The same number of labels were to be used on each of the paper and specimen pairs of tests, except that in the case of the gladiolus diagram three sepals and two petals had to be labeled correctly in order to get one point for each. the large transverse wood sections had to be polished to show the structure.

ADMINISTRATION OF TESTS

The diagram test was given in one period. The students were asked to go as rapidly as they conveniently could but were told that they might take what time they needed. The practical test was given two days later without warning. The students did not know that a second test was coming. Each student doing the practical test was given a sheet of clean cardboard on which was placed a clam, gladiolus flower, and a polished cross-section of Lime or Linden stem. Each was given the mimeographed labels, and sufficient pins to attach labels. Students were given scalpels, probes and scissors, so that the parts to be labelled could be dissected or left entire whichever they preferred. Labels could be attached any way provided they were on the correct organs. On completing his labelling the student had to place his named cardboard on a side table for marking. At the end of the period (just before lunch) three trained biology students marked the results according to definite instructions. As these students themselves were able and knew this work, the corrections would likely be quite accurate. This was borne out when the investigator checked at random five sets. These were found to be marked correctly. The students also corrected the paper tests from prepared keys. Scores on each organism were kept separate and correlations "paper vs. practical" run for each set.

PLACE THE PROPER TERM AS LABEL IN THE BLANK SPACES PROVIDED:
Cortex, foot, sepal, mouth, pith, xylem, sepal, heart, anther, ovary, stomach and intestinal mass, style, bract, phloem, adductor muscles, cambium, gills, stigma, mantle, petal, retractor muscle, medullary ray, siphon, filament, bark, pallial muscle, shell.

... 1 ...
... 2 ...
... 3 ...
... 4 ...
... 5 ...
... 6 ...
... 7 ...
... 8 ...
... 9 ...
... 10 ...
... 11 ...)
Results. The paper test of diagrams took fourteen minutes for the last student to complete while the specimen test took thirty-five minutes for the last to complete the test.

Only correct labellings were counted. No deductions were made for errors.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B,J</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>C,G</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>C,L</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>F,P</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F,A</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>F,M</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>G,V</td>
<td>8</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>G,I</td>
<td>9</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>H,B</td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>J,D</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>K,T</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>K,M</td>
<td>5</td>
<td>11</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>L,J</td>
<td>4</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>L,D</td>
<td>11</td>
<td>10</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>M,D</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>M,H</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>M,J</td>
<td>5</td>
<td>10</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>N,B</td>
<td>2</td>
<td>8</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>P,I</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>P,J</td>
<td>5</td>
<td>10</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>P,D</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>R,F</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>S,F</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>S,K</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>S,S</td>
<td>11</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>S,B</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>S,T</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>S,K</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>S,R</td>
<td>3</td>
<td>10</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>W,G</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>W,A</td>
<td>5</td>
<td>9</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>W,G</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>4</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>W,B</td>
<td>11</td>
<td>10</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

Ave. 6.65 8 5.97 6.06 5.74 4.23

Correlations 0.56 ± 0.07 0.76 ± 0.04 0.54 ± 0.08
Findings  Judging by the average marks made by 34 students on the Clam and the Flower test, students make better marks on the practical than on the paper diagram test.

<table>
<thead>
<tr>
<th></th>
<th>Clam</th>
<th>Flower</th>
<th>Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPER</td>
<td>6.65</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>PRACTICAL</td>
<td>5.97</td>
<td>6.06</td>
<td>4.1</td>
</tr>
</tbody>
</table>

The investigator is of the opinion that this not due to learning effect. The Stem test showed the reverse of the first two to add support to this statement. No corrections were made. No books were borrowed from the room library in the interim.

The reason for the lower marks in the Stem test came out in the marking of the tests. Because the cambium, phloem, and cortex in the actual specimen occupied a ring about one to two millimetres wide the mechanical difficulties of differentiation were great. This was borne out in asking certain students after what difficulties they had met (Students did know their scores). All reported the same difficulty about the size of the ring. (It seems obvious that if other teachers wished to do testing of this type they must make sure that the specimens are large enough to label all parts asked, or else use microscope examination techniques when dealing with material of this type.)

The correlations were fairly high for the first two, .66±.07 and .78±.04 while only .54±.08 for the stem. On the various tests most made (identical pairs of marks on the Flower, 13 against 6 pairs of identical marks for the Clam and none for the Stem.

<table>
<thead>
<tr>
<th></th>
<th>Clam</th>
<th>Flower</th>
<th>Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identical scores</td>
<td>6</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Higher scores on practical test</td>
<td>19</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>Higher score on paper test</td>
<td>9</td>
<td>8</td>
<td>33</td>
</tr>
</tbody>
</table>

Seventeen out of thirty-four students did better to the extent of two points out of nineteen on the diagram test for the first two (Clam and Flower). The stem test was not included for a special problem exists there (as shown later). Twelve did approximately equally, to a tolerance of one point. Five did better on the diagram test. (Four of these are known to the investigators to be "text memorizers" preferring not to work with the specimens provided. These results suggest also another problem, that of the relative worth and time-worth of the two methods of presenting material and testing.

It is necessary to note here that the students at all times have specimens or materials to examine in all cases, where these would be of any service. These results conceivably might be at variance with those obtained where classes are taught largely from text books or diagrams only without specimens.

For the time spent a paper test will be more economical but shows only a correlation of .66 with the actual specimen testing, but because it is generally the desire of most progressive educators to get away from the abstract to the concrete wherever possible the results of this test would suggest that more testing of a practical nature should be incorporated into any testing programme.
Questions involving the handling of actual things need to be prepared with the utmost thought for practical difficulties, such as manipulation and size of objects.

Above all it must be realized that although the correlations are positive and moderate, still they are far too low for a teacher to use a paper test as an equivalent or substitute for a "practical test".
PURPOSES

(a) To check the results of the third portion of the preceding test.
(b) To find what correlation exists between the ability to identify the tissues of a transverse section (stained and prepared) of Tilia stem (a woody stem) and the ability to label a reasonably accurate diagram of the same kind of stem used in the preceding experiment.

PROCEDURE Because the results of the third section of the preceding test were so much at variance with the results of the first two parts the investigator suspected that the two forms of the tests were not as equal in difficulty as they might be. Largely by inspection of pupils at work labelling the polished transverse stem sections of wood the investigator could see great difficulties encountered in separating these three tissues: cambium, phloem, cortex. The area of these three combined was not very great, and differentiation of tissues was difficult.

To bring the practical side of the test in line with the diagram used in the preceding experiment the microscope slide of Tilia stem from which the diagram had been taken was put under a microscope having a demonstration eyepiece and indicator needle. The section was stained with methylene blue and safranin and magnified sixty (60) diameters. All the names of the tissues were put on the blackboard so that each student would be familiar with the terms used.

No student had been told his score on the preceding test, nor had he corrected a paper, nor seen his own corrected paper. No student knew that the second test was to be presented so that on the basis of chance and human inertia there would be very little further study on this topic, in the interval between tests, especially as the students had been notified that this was an experiment and they had been asked to carry out instructions carefully.

Each student was summoned individually to the microscope where the investigator took the main eyepiece and needle indicator and the student the branch eyepiece. First the microscope was focused so that the student had the clearest field possible. (Sometimes focal lengths of person's eyes differ). Then a survey of the entire stem was presented to the student without comments. This was at a speed controlled by the student's request. Then the student was asked to name the various parts of the stem. (In order that no memorization of a sequence of answers was possible by those who had not yet been summoned to the microscope the order of answers for each individual was different. Correct answers only were counted. If the student were hesitant the question was repeated. Wrong answers were corrected there.

The rest of the class worked at an unrelated assignment.

RESULTS

This microscope test took forty-one minutes for 34 students taken individually.

The scores were:
<table>
<thead>
<tr>
<th>STUDENT</th>
<th>SCORE ON DIAGRAM TEST</th>
<th>MICROSCOPIC TEST SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>B,J</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>B,J</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>C,G</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>C,L</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>F, P</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>F, A</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>F, M</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>G, V</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>G, I</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>H, B</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>J, D</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>K, T</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>K, M</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>L, J</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>L, D</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>M, D</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>M, H</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>McL J</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>N, B</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>P, I</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>P, J</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>F, D</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>R, F</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>S, G</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>S, K</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>S, S</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>S, B</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>S, T</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>S, K</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>S, K</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>W, G</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>W, A</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>W, G</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>W, B</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

**Total** 195 193

**Average** 5.74 5.67

**Correlation** \( .69 \pm .06 \)
FINDINGS  It is obvious that this pair of tests is more nearly parallel than the first pair. The practical test yield higher scores than in the previous method. The correlation was about the same as in the first two portions of the second experiment, namely .69 ± .06.

EXPLANATION: A possible explanation of the higher scores and correlation would lie in the fact that the cambium-phloem-cortex tissues were magnified more nearly to the size of the diagram. Then again the diagram used was taken from a microscopic view.

The results here together with the results of the preceding experiment should make it clear that practical testing material must be chosen with great care. From these experiments it would appear wise for a teacher to try out any practical test on a small group first.
EXPERIMENT IV

PURPOSES (a) To see in which form of test, actual manipulation or a "paper test", based on simple laboratory exercises the pupils of high school classes show better results.

(b) To find the correlation between marks of each type of test.

PROCEDURE Only two classes were available for testing namely a Grade XI General Science V, and a Grade X General Science. The Grade IX class available had no training in these laboratory techniques so that they would be of little value.

For the purpose of this experiment two common laboratory activities were chosen:
(a) the measuring of volumes of liquids
(b) the weighing of articles.

Each of these two activities will be treated separately here.

(a) Four graduated vessels, 250 C.C., 100 C.C., and two 10 C.C. graduate were chosen. In the largest one 162 C.C. of water were placed, in the second 66 C.C. of methyl hydrate (C\textsubscript{2}H\textsubscript{6}OH), were added, in the third one 3.2 C.C. of concentrated sulphuric acid, and in the last one 6.2 C.C. of mercury.

A series of four diagrams was prepared with the level of the water shown by lines. The meniscus curve was shown. The levels shown of the liquids were different from the four actual cases in order to avoid a direct transfer. (See page 175.)

In testing the students the paper test was given in the usual way: when finished the papers were collected, and marked by the examiner only. Because only about fifteen minutes of the class period had been used to do this test on paper the first part of the practical test, (volume measuring) was given. The four graduated vessels were numbered and placed on tables in the front of the room. Each student was given a blank sheet of paper on which he had to write his name and class, and then number spaces 1-4. The students were arranged in a single file and brought in front of the graduate on the north side of the room first, passing southward to the next, and so forth. At the south end of the line the student's paper was collected so that in the arrangement practically no chance of copying occurred.

RESULTS

The two sets of scores were tabulated.

(over)
<table>
<thead>
<tr>
<th>GENERAL SCIENCE V</th>
<th>Actual</th>
<th>Measuring</th>
<th>Diag.</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A, T</td>
<td>2</td>
<td>3</td>
<td>A, A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A, E</td>
<td>1</td>
<td>0</td>
<td>B, H</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>A, N</td>
<td>2</td>
<td>3</td>
<td>B, A</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>B, J</td>
<td>2</td>
<td>0</td>
<td>B, D</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B, D</td>
<td>2</td>
<td>0</td>
<td>C, J</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>C, P</td>
<td>2</td>
<td>0</td>
<td>C, M</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>E, H</td>
<td>0</td>
<td>1</td>
<td>C, M</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>F, F</td>
<td>2</td>
<td>4</td>
<td>C, R</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>G, Lino</td>
<td>1</td>
<td>0</td>
<td>D, M</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>G, Law</td>
<td>2</td>
<td>0</td>
<td>D, R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>G, A</td>
<td>3</td>
<td>2</td>
<td>H, S</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>G, P</td>
<td>1</td>
<td>1</td>
<td>H, A</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>H, W</td>
<td>1</td>
<td>2</td>
<td>K, C</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>H, H</td>
<td>2</td>
<td>0</td>
<td>K, T</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>H, R</td>
<td>0</td>
<td>1</td>
<td>K, W</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>H, Wm</td>
<td>2</td>
<td>4</td>
<td>L, V</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>K, J</td>
<td>2</td>
<td>0</td>
<td>K, H</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>L, M</td>
<td>3</td>
<td>2</td>
<td>KPB, F</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>L, V</td>
<td>1</td>
<td>0</td>
<td>L, E</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>L, D</td>
<td>2</td>
<td>3</td>
<td>P, E</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>M, N</td>
<td>2</td>
<td>2</td>
<td>R, W</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>M, C</td>
<td>2</td>
<td>3</td>
<td>R, M</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>M, I</td>
<td>0</td>
<td>3</td>
<td>R, T</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>M, Y</td>
<td>2</td>
<td>1</td>
<td>R, J</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>M, R</td>
<td>2</td>
<td>1</td>
<td>R, H</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>P, M</td>
<td>2</td>
<td>1</td>
<td>R, M</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S, F</td>
<td>2</td>
<td>3</td>
<td>R, M</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>S, A</td>
<td>2</td>
<td>1</td>
<td>S, J</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>S, L</td>
<td>2</td>
<td>4</td>
<td>S, J</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>T, S</td>
<td>1</td>
<td>0</td>
<td>S, E</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>W, S</td>
<td>2</td>
<td>0</td>
<td>S, E</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>W, H</td>
<td>1</td>
<td>2</td>
<td>S, G</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W, T</td>
<td>3</td>
<td>2</td>
<td>T, M</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W, J</td>
<td>2</td>
<td>3</td>
<td>T, P</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W, E</td>
<td>3</td>
<td>2</td>
<td>W, F</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Z, E</td>
<td>1</td>
<td>2</td>
<td>L, R</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>S, M</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL**       | 63     | 63        |       | 39      | 40      |
**AVERAGE**     | 1.7    | 1.7       |       | 1.08    | 1.11    |
**CORRELATION** | .24 ± .12 |         |       | .24 ± .1 |         |
FINDINGS From the average scores it might be deduced that there is an acquisition of manipulatory skills as a result of General Science IV work as conducted in Britannia High School. As can be seen from an examination of the calculations the correlations were very low (-.24 ± .12 and .24 ± .10) showing that the two methods of testing are not likely testing the same thing. Teachers evidently would be ill advised to consider a paper test of an actual manipulating skill as equivalent to a direct test.
PROCEDURE A paper test was prepared explaining what weights were used, where articles were placed on balance, and such (as shown on the sample test paper).

A beam balance of a good grade was provided for each pair of students. A series of pebbles was weighed by the instructor, numbered, and the weights recorded. Each student was given a pebble told to weigh it as accurately as he could by himself, record the weight on a slip of paper, sign his paper, and turn the paper in. Having each pebble different insured that no copying could be done in a rather crowded laboratory.

SAMPLE OF TEST

RESULTS ON NEXT PAGE

FINDINGS

Again the correlations are very low, (.21 ± .10 and .58 ± .08 respectively). The two types of tests could not be considered equivalent as they obviously are not testing the same abilities. Actual skills must be tested directly rather than by means of paper tests. One class, Class 10, shows a much higher correlation on this than the other class does. So far the investigator has no explanation. Even the higher correlation, which approximates those in Experiment 111, is not high enough to warrant the use of a paper test to examine a skill, or any manipulative test.
<table>
<thead>
<tr>
<th>Student</th>
<th>Actual Weighing</th>
<th>Diagram</th>
<th>Student</th>
<th>Actual Weighing</th>
<th>Diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A</td>
<td>2</td>
<td>0</td>
<td>A, A</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>A, E</td>
<td>2</td>
<td>0</td>
<td>B, H</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>A, N</td>
<td>2</td>
<td>0</td>
<td>B, A</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>B, J</td>
<td>2</td>
<td>0</td>
<td>B, D</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>B, D</td>
<td>2</td>
<td>0</td>
<td>C, J</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C, F</td>
<td>2</td>
<td>0</td>
<td>C, M</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F, H</td>
<td>2</td>
<td>0</td>
<td>C, R</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F, P</td>
<td>0</td>
<td>0</td>
<td>D, M</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>G, L</td>
<td>1</td>
<td>0</td>
<td>D, K</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>G, A</td>
<td>2</td>
<td>0</td>
<td>D, S</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>G, P</td>
<td>3</td>
<td>1</td>
<td>J, C</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>H, W</td>
<td>1</td>
<td>0</td>
<td>K, T</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>H, H</td>
<td>1</td>
<td>0</td>
<td>K, W</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>H, R</td>
<td>1</td>
<td>0</td>
<td>L, V</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>H, W</td>
<td>2</td>
<td>0</td>
<td>M, H</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>J, K</td>
<td>2</td>
<td>0</td>
<td>M, P</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>J, L</td>
<td>2</td>
<td>1</td>
<td>P, E</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>L, D</td>
<td>2</td>
<td>0</td>
<td>P, W</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>M, N</td>
<td>1</td>
<td>0</td>
<td>P, M</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>M, C</td>
<td>2</td>
<td>0</td>
<td>R, T</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>M, B</td>
<td>3</td>
<td>0</td>
<td>R, J</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>M, Y</td>
<td>2</td>
<td>0</td>
<td>R, H</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>M, R</td>
<td>2</td>
<td>1</td>
<td>R, M</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P, M</td>
<td>2</td>
<td>0</td>
<td>R, W</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>S, F</td>
<td>3</td>
<td>0</td>
<td>S, J</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>S, A</td>
<td>2</td>
<td>0</td>
<td>S, J</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>S, L</td>
<td>3</td>
<td>0</td>
<td>S, K</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>T, S</td>
<td>3</td>
<td>1</td>
<td>S, E</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>W, S</td>
<td>3</td>
<td>0</td>
<td>S, G</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W, H</td>
<td>0</td>
<td>0</td>
<td>T, M</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>W, T</td>
<td>0</td>
<td>0</td>
<td>T, P</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>W, J</td>
<td>0</td>
<td>0</td>
<td>W, F</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>W, E</td>
<td>0</td>
<td>0</td>
<td>L, R</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**TOTAL** 66 39  **TOTAL** 53 30

**AVERAGE** 1.71 1.05  **AVERAGE** 1.47 0.83

**CORRELATIONS** 0.21 1.10  **CORRELATIONS** 0.58 0.07
EXPERIMENT V

PURPOSE OF TEST:
(a) To see if there exists an increase from grade to grade in the amount of reading of scientific books in the school library.
(b) To see if the test and technique measures this suspected increase (the validity of test)

PROCEDURE
Preparation of Test
First, a complete survey of all magazines and books in the school library was made. Each magazine or book was read or sampled by the examiner in order to become familiar with the contents. Then questions of a broad nature based on major topics or characteristics of the book or magazine were prepared together with spurious answers that might be deduced from the titles in order to check those students who might be tempted to guess the answers.

Because there were not very many magazines with scientific material in them a matching-question group was feasible. For the books this form seemed unsuitably, mainly because of the number of books, and the rarity of the use of some. A modification of the true-false type seemed most easily prepared and the most compact form of question. Multiple choice would demand much more time to prepare, and for the present was avoided. Completion or recall questions are almost an impossibility for this work due either to the tremendous number of correct responses to broad or general questions that are possible, or to the too precise or minute type of question that evolves when questions are prepared in a way to reduce the great number of possible correct responses to the other type. A series of "yes-no" questions in connection with the titles of all the books materialized from the various attempts to prepare a test.

One rather important difference between preparing an ordinary information test or similar type, and questions to find intent of reading is that of sampling. For a survey of the reading done by students all titles of books in the particular library section must be included. The only sampling permitted would occur in the selection of statements about the book.

The test was mimeographed. A copy follows.

(b) Administration of Test
The test was given to the same classes as in the preceding tests, namely,
Class 17 Grade IX 1938-39 Britannia High School
Class 10 Grade X " " "
General Science V Grade XI 1938-39 Britannia High School

The same limitations of numbers and lack of equating classes will be present in the results of the experiment.

To use the Library Loan Cards and records for the source of material for this experiment was not considered sound because:
1. The school library had been changed from an antiquated method of handling, cataloguing, and recording to the Dewey decimal system and loan cards after Christmas of the year 1937. There are no records preceding this date.
2. Many students read books in spare periods, library periods, noon hours and after school right in the library leaving no records to check. There are reserve shelves with no records kept.
Administering the Test (cont.) The class was informed of the purpose of the test and put at ease regarding low scores likely to result. They were asked not to guess but to answer only those questions which concerned books that they read. These directions were felt to be necessary in order to reduce to a minimum the desire of some students to get a high score, for it doubtless would be possible to obtain a higher net score on this test by some shrewd guessing. As a further deterrent against deliberate guessing the students were told that their scores would not be given to them after the correction of the test, as it was felt that this action would tend to eliminate rivalry for high scores.

The class was informed that it could have all the time that any particular person needed. Also the class was informed that it would take about twenty-five to thirty minutes to answer the test. (Some students actually took the full forty minutes.) Of course rigid control over the group to prevent copying or communication of information was exercised. The scores of two pupils who were in collusion were deleted from the records so that the examiner felt that true net scores had been obtained. In any event their gross scores were lower than the class average.

The directions on the test stated that "wrongs" would be subtracted from the "correct" gross score. The examiner as a result of correcting the paper feels that this procedure was justified for on several papers entries were made against several books that the examiner suspected had not been read. The corrections made by calculation on sixteen papers (checked by later personal interviews) showed net scores of zero, which coincided with admissions from these sixteen that they had not read any of the books.

TEST OF USE OF SCIENCE BOOKS IN THE LIBRARY

1. Place the letter that is opposite the titles of magazines found in our library in the blank in front of the statement that describes that magazine, or applies to it. Two encyclopaedias have been included in this list also, and in questions.

a. Punch
b. C I L Oval
c. Journal of Chemical Education
d. Canadian Geographical Journal
e. L'Illustration
f. Scientific American
g. National Geographic Magazine
h. Modern Mechanic
i. Nature Magazine
j. World Book
k. London Illustrated News
l. Encyclopaedia Britannica
m. Popular Mechanics

1. .................Contains rather technical articles concerning the investigations into the composition of substances; is fairly up-to-date.
2. .................good source of geological material for the northern part of the continent by virtue of its excellent photographs of geological structures.
3. contains a wealth of plans and directions for constructing many kinds of things; also a few articles of semi-scientific nature.
4. a magazine that gives an up-to-date report on many fields of science, and in a way that can be understood by the average high school pupil.
5. frequently contains many photographs, diagrams, articles on the development of airplanes and aviation, particularly warplanes of the world.
6. suitable for a moderately precise exposition of some of the older contributions in all fields of science.
7. reports the activities of one of the largest commercial chemical manufacturers in Canada.
8. very frequently produces excellent articles, colored plates, etc. on the geology of almost any part of the world.
9. contains many articles, stories, etc., of a scientific nature, and written in a style that makes it very suitable for junior high and similar grades.
10. often contains excellent articles of a natural history type, with first-class plates.
11. contains a page called "The World of Science" written by foremost scientists, and usually biological in nature; often there are other first-class articles on anthropology and biology.

11. The following exercise is to see how familiar you are with the science books of our library. There is a title of a book, together with the author's name, and one or more statements about the book. If, as a result of your reading, you recognize that the book mentioned contains material of the type described place a "YES" in the blank in front of this statement; but if you think that the book does not contain this kind of information place a "NO" in the blank. If you know nothing about the book leave the space blank; do not guess as the errors will be deducted from the other marks.

BIOLOGY FOR TODAY; Caldwell and Curtis.
1. an extremely technical and difficult text explaining accurately the functions that are common to all forms of life.
2. based mainly on the idea that all organisms are continually struggling to secure energy in order to carry on other activities.

THE NATURE AND DEVELOPMENT OF PLANTS; Curtis
3. A book explaining the nature of plants and flowers in a way satisfactory for students from grades seven to ten.

TREES AND SHRUBS OF BRITISH COLUMBIA; Anderson
4. Contains many photographs, diagrams, and chatty articles about our native trees.

MARVELS OF FISH LIFE; Ward
5. A very readable book explaining reports of such phenomena as changing of color, and protective coloration of fish.
THE TEST (CONT'D)

ILLUSTRATED FLORA OF THE NORTHERN STATES AND CANADA; Britton & Brown
6.....................A series of three volumes of books with line diagrams and sketches of the great majority of plants found in our region.

BIRDS OF AMERICA; Nature Lovers Library. Volume 1,11,111
7.....................A catalogue only of the birds found in America.

MAMMALS OF AMERICA. Nature Lovers Library. Volume IV.
8.....................Contains articles and colored plates, etc., about frogs, fish, and snakes of America.

DEVILS, DRUGS, AND DOCTORS. Haggard.
9.....................Contains chapters telling of the various ways that mankind has cared for or neglected mothers and babies.
10.....................Gives a comparison of the methods used by the "doctors" and medicine men of Egypt, India, China, The Indians of Americas, and many backward people.
11.....................Contains chapters on "Pestilences and Moralists", and the social diseases.
12.....................Criticises our modern doctors as being selfish and not well trained.

MAN VS. MICROBES; Kopeloff.
13..............Relates the story of bacteriology, its rise and present importance.
14.....................Certain microbes are "harnessed to industry", soil bacteria are useful.
15.....................Contains chapters on intestinal bacteria, lockjaw, botulism, cholera, yellow fever.
16.....................Shows how smallpox, colds, and infantile paralysis (poliomyelitis) are all bacterial diseases.

FODDER AND PASTURE PLANTS. Clark and Malte
17.....................Illustrated book to assist in identifying, and to spread information and advice about, the chief plants used for livestock.

AUDEL'S GARDENERS' AND GROWERS' GUIDE
18.....................Explains where and how to market produce; Market boards "pool, investigation".
19.....................Gives many useful hints to the average city gardener.

TEXTILE FIBRES. Mathews
20.....................Explains how the chief clothing materials of the world are produced.

HOW TO KNOW TEXTILES. Henley
21.....................Gives ready means of analyzing various fabrics to find actual components; the uses of the several fabrics.

TWENTIETH CENTURY BOOK OF RECIPES. Henley
22.....................An excellent book that goes into all phases of cooking and cookery. Should be read by all future cooks.
THE TEST (CONT'D).

MAKERS OF SCIENCE.
23. ................. A book that gives biographies of the chief scientists in each of the fields of physics, chemistry, biology, and astronomy.

THE RADIO AMATEUR'S HANDBOOK. Collins
24. .................. Goes into the theoretical calculations and laws concerning the use of radiant energy as a carrier of our messages.

FLYING AND HOW TO DO IT. Jordanoff.
25. .................. A very chatty simple book explaining the first principles of flying.
26. .................. No diagrams or plates is a serious fault in the book.

OUTLINE OF SCIENCE. Sir Jas. Thomson
27. .................. A masterly series of books that deals in the main with the developments in chemistry and physics.

THE UNIVERSE AROUND US. Sir Jas. Jeans
28. .................. "Exploring in Time", "Beginnings and Endings".
29. .................. "Exploring the atom"

THE THEORY OF RELATIVITY. Einstein.
30. .................. A very easy explanation of how there is no absolute time or velocity by the man who first developed this idea.

THE NEW KNOWLEDGE. Duncan.

DISCOVERY, THE SPIRIT AND SERVICE OF SCIENCE.
32. .................. Shows how scientists are really explorers into the realms of the unknown.

ELEMENTS OF PHYSICS. Smith
33. .................. A very interesting and easily read book on the rudiments of physical forces, of the average high school student's level.

PHYSICS FOR COLLEGE STUDENTS. Knowlton.
34. .................. A book dealing with dynamics, light, heat, sound, electricity, and magnetism in the advanced manner.

EXPERIMENTAL SCIENCE. Collins.
35. .................. Contains directions and many explanations for simple experiments that can be performed at home or school, to illustrate certain truths.

SOUND. Richardson
36. .................. A simple book devoted to an easy explanation of musical and harmony.

CREATIVE CHEMISTRY. Slosson.
37. .................. Contains chapters on high explosives and gases used in warfare.
38. .................. Explains how perfumes, and dyes, and drugs are made.
CREATIVE CHEMISTRY. Slosson.
39. Goes into great details concerning the structure of atoms.

CHEMISTRY AND CIVILIZATION. Gushman.
40. Shows how certain discoveries in chemistry have changed drastically the civilization in which we now find ourselves.

THE DISCOVERY OF ELEMENTS. Weeks.
41. A book written by a woman who tells in a very interesting manner the toils and troubles, of the brilliant successes of chemists in searching for new elements.

THE CARBON COMPOUNDS. Porter.
42. Deals with the organic compounds telling how they are derived; benzene ring.

THREE CENTURIES OF CHEMISTRY. Masson
43. The great contributions of chemists from the days of the Moors to the present time.

UNDER THE TROPIC SEAS. Wm. Beebe
44. Explains how treasure is recovered from the hulks of sunken Spanish galleons in and around the Caribbean Sea.

THE PET BOOK. Comstock.
45. Explains how to care for pets.
46. Relates of the other peoples of the earth and their pets.

SMITHSONIAN SCIENTIFIC SERIES.
47. Has volumes on great inventions, prehistoric man, insects.
48. Deals in another volume with minerals from earth and sky.
49. Contains a volume of cold-blooded animals; frogs, fish and snakes.

THE OLD RED SANDSTONE. Miller
50. A recounting of the discovery and an explanation of the importance of this deep layer under England.

THE ORIGIN OF THE EARTH. Chamberlain
51. Supports the hypothesis that the solar system has condensed from a rotating mass of gases.

WATER INTO GOLD. E. Hill.
52. Paints a vivid picture of attempts to reclaim the tremendous amounts of colloidal gold in sea water.
53. Describes the raising of grapes and raisins in Australia.

MORE FOR YOUR MONEY. Bennett.
54. Tells you how to bargain with any person who is attempting to sell you something, particularly how to guard against high pressure salesmen.
THE TEST (CONT'D).

RATS, LICE AND HISTORY. Zinsser.
56....................Describes graphically the cause of bubonic plague
(the Black Death), its prevention, trench fever.

THE NATURE OF THE PHYSICAL WORLD.
56....................Rather philosophical analysis of the world of physics
as we know it now.

100,000,000 MILLION GUINEA PIGS. Kallet and Schlink.
57....................Speaks of the American public as experimental pawns
of the many get-rich-quick schemes in patent medicines.
58....................Tells what widely advertised drugs and patent medicin­
es, antiseptics, etc. are worthless or worse.

FORESTS OF BRITISH COLUMBIA. Whitford and Craig.
59....................A technical account of the distribution and value of
the various forest trees of B. C.
## RESULTS

### Britannia High School

**Class 17 1938 (Sept.)**

<table>
<thead>
<tr>
<th>Student</th>
<th>Scores</th>
<th>Scores</th>
<th>Scores</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Magazines</td>
<td>Books</td>
<td>Magazines</td>
<td>Books</td>
</tr>
<tr>
<td>A, J</td>
<td>0</td>
<td>2</td>
<td>F, T</td>
<td>6</td>
</tr>
<tr>
<td>C, M</td>
<td>0</td>
<td>0</td>
<td>F, N</td>
<td>3</td>
</tr>
<tr>
<td>C, L</td>
<td>7</td>
<td>0</td>
<td>R, M</td>
<td>3</td>
</tr>
<tr>
<td>D, D</td>
<td>4</td>
<td>0</td>
<td>S, M</td>
<td>2</td>
</tr>
<tr>
<td>L, I</td>
<td>2</td>
<td>4</td>
<td>S, B</td>
<td>2</td>
</tr>
<tr>
<td>L, F</td>
<td>0</td>
<td>3</td>
<td>S, C</td>
<td>0</td>
</tr>
<tr>
<td>L, B</td>
<td>0</td>
<td>0</td>
<td>S, D</td>
<td>0</td>
</tr>
<tr>
<td>L, S</td>
<td>2</td>
<td>5</td>
<td>S, A</td>
<td>0</td>
</tr>
<tr>
<td>M, H</td>
<td>3</td>
<td>0</td>
<td>T, H</td>
<td>0</td>
</tr>
<tr>
<td>M, H</td>
<td>2</td>
<td>0</td>
<td>T, Y</td>
<td>0</td>
</tr>
<tr>
<td>M, R</td>
<td>2</td>
<td>2</td>
<td>T, M</td>
<td>0</td>
</tr>
<tr>
<td>M, S</td>
<td>3</td>
<td>0</td>
<td>U, K</td>
<td>3</td>
</tr>
<tr>
<td>McA, F</td>
<td>1</td>
<td>0</td>
<td>W, T</td>
<td>3</td>
</tr>
<tr>
<td>McD, S</td>
<td>1</td>
<td>0</td>
<td>W, G</td>
<td>2</td>
</tr>
<tr>
<td>McG, J</td>
<td>0</td>
<td>1</td>
<td>W, D</td>
<td>5</td>
</tr>
<tr>
<td>McI, H</td>
<td>3</td>
<td>0</td>
<td>W, J</td>
<td>2</td>
</tr>
<tr>
<td>McK, D</td>
<td>2</td>
<td>0</td>
<td>W, E</td>
<td>1</td>
</tr>
<tr>
<td>McW, A</td>
<td>5</td>
<td>1</td>
<td>W, J</td>
<td>3</td>
</tr>
<tr>
<td>N, M</td>
<td>2</td>
<td>0</td>
<td>Y, T</td>
<td>0</td>
</tr>
<tr>
<td>N, E</td>
<td>3</td>
<td>1</td>
<td>TOTAL</td>
<td>84</td>
</tr>
</tbody>
</table>

**AVERAGE**

2.15  .51
<table>
<thead>
<tr>
<th>STUDENT</th>
<th>MAGAZINES</th>
<th>BOOKS</th>
<th>STUDENT</th>
<th>MAGAZINES</th>
<th>BOOKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, A</td>
<td>5</td>
<td>0</td>
<td>M, H</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>B, H</td>
<td>5</td>
<td>2</td>
<td>McB, F</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>B, A</td>
<td>3</td>
<td>4</td>
<td>McV, E</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>B, D</td>
<td>3</td>
<td>3</td>
<td>P, E</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>C, J</td>
<td>5</td>
<td>6</td>
<td>P, W</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>C, M</td>
<td>6</td>
<td>0</td>
<td>P, M</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>C, M</td>
<td>6</td>
<td>0</td>
<td>R, T</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>C, R</td>
<td>2</td>
<td>9</td>
<td>R, J</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D, M</td>
<td>6</td>
<td>4</td>
<td>R, H</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>D, R</td>
<td>6</td>
<td>0</td>
<td>R, M</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>F, D</td>
<td>6</td>
<td>0</td>
<td>R, M</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>G, A</td>
<td>5</td>
<td>1</td>
<td>S, J</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>H, S</td>
<td>7</td>
<td>0</td>
<td>S, J</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>H, A</td>
<td>4</td>
<td>3</td>
<td>S, E</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>J, C</td>
<td>5</td>
<td>1</td>
<td>S, N</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>K, T</td>
<td>5</td>
<td>6</td>
<td>S, G</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>K, W</td>
<td>5</td>
<td>2</td>
<td>T, M</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>L, Y</td>
<td>3</td>
<td>4</td>
<td>T, F</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>M, N</td>
<td>0</td>
<td>0</td>
<td>W, F</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

TOTAL 156
AVERAGE 2.6
<table>
<thead>
<tr>
<th>STUDENT</th>
<th>MAGAZINE</th>
<th>BOOK</th>
<th>STUDENT</th>
<th>MAGAZINE</th>
<th>BOOK</th>
</tr>
</thead>
<tbody>
<tr>
<td>A, T</td>
<td>4</td>
<td>4</td>
<td>L, V</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>A, EJ</td>
<td>4</td>
<td>2</td>
<td>L, P</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>A, N</td>
<td>5</td>
<td>5</td>
<td>M, N</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>B, J</td>
<td>3</td>
<td>6</td>
<td>M, C</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>B, D</td>
<td>8</td>
<td>2</td>
<td>M, I</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>C, Y</td>
<td>4</td>
<td>2</td>
<td>M, Y</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>C, P</td>
<td>2</td>
<td>3</td>
<td>McC, E</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>E, H</td>
<td>7</td>
<td>3</td>
<td>P, M</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>E, F</td>
<td>3</td>
<td>4</td>
<td>S, F</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>G, L</td>
<td>3</td>
<td>3</td>
<td>S, M</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>G, L</td>
<td>6</td>
<td>1</td>
<td>S, A</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>G, A</td>
<td>4</td>
<td>5</td>
<td>S, L</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>G, P</td>
<td>3</td>
<td>0</td>
<td>T, S</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>H, W</td>
<td>5</td>
<td>0</td>
<td>W, S</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>H, H</td>
<td>3</td>
<td>2</td>
<td>W, H</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>H, R</td>
<td>3</td>
<td>3</td>
<td>W, T</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>H, W</td>
<td>5</td>
<td>10</td>
<td>W, J</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>K, J</td>
<td>7</td>
<td>1</td>
<td>W, E</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>L, M</td>
<td>3</td>
<td>3</td>
<td>Z, E</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>168</strong></td>
<td></td>
<td><strong>134</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td><strong>4.15</strong></td>
<td></td>
<td><strong>3.53</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FINDINGS  Judging by the average scores from the various classes it seems that the test does actually measure the students' use of the science library. There is a yearly increment of books read from a Grade IX average score of .51 points (not books), through a Grade X average of 2.6 points, to a Grade XI average of 3.55 points. (Fifty-nine points covered forty-two books).

It seems that in the school tested that familiarity with magazine content increases materially in Grade IX. Newcomers to the school are well acquainted with Popular Mechanics, and the National Geographic, but during the year become familiar with half a dozen other magazines. Scientific American and the Journal of Chemical Education are scarcely touched by any students up to Grade XI. Any use of these if at all must come in Grade XII which was not tested at all.

The test had an excellent teaching value for it drew the attention of many students to these books and magazines in the library. Unfortunately the investigator did not anticipate this so no preparations were made to correlate loans, readings, in order to compare these results with previous conditions. The examiner was asked many more questions about the books, and more students brought books to him to discuss certain articles than occurred at any other time since the establishment of the library.

All in all, the test seems to show fair promise of a reasonably objective method to secure data on the extent of the use of the science section of the library. Quite likely the technique could be transferred into any other department. Doubtless the test could be improved with further experience. It would need to be enlarged each year because of additions to the library. The technique would likely break down when used for "extensive scientific libraries" as in Universities, where possibly a division of the fields might be wiser. Another point that might be gleaned from the experiment is the answer to How Many Books Should an Average Student Read Each Year? Because a score of .51 represents (.51 times 59) or .37 approx. of a book which the Grade IX students seem to have read in one way or another before they entered high school, it would seem that at the end of the first year they had read (2.6 times 42) minus .37) or 1.48 (approx.) of a book. This amount would suggest a C value if any credit were given to students for books read, but because the results were obtained from only one class (average ability) further need of testing is advisable before this figure should be taken as a guide. During the Grade X year (3.52 times 42) minus (.51 plus 1.48) or .65 of a book is read. This figure is possibly quite away from the class tested is notoriously poor in application, though average in ability. Nevertheless, the method of obtaining much wanted data seems clear, needing only more cases to settle the point. That is, classes in each grade should be given the same test at the same time, late in June or early in September. In order to secure the average, increments of preceding years must be deducted from the average score for that grade.