1949

The construction and evaluation of a general science inventory test for grade nine (based on vocabulary).

BOSTON UNIVERSITY

SCHOOL OF EDUCATION

Thesis

THE CONSTRUCTION AND EVALUATION OF A GENERAL SCIENCE INVENTORY TEST FOR GRADE NINE

(based on vocabulary)

Submitted by

Mildred F. Emerson

(B. Ed., Keene Teachers' College, 1938)

In partial fulfillment of the requirements

for the Degree of Master of Education

1949

School of education
June, 1949
30738

Thesis approved:

First Reader:	John G. Read Associate Professor of Science Education
Second Reader:	Henry W. Syer Assistant Professor of Education
Third Reader:	William C. Kvaraceus Associate Professor of Education

ACKNOWLEDGMENT

Grateful appreciation is expressed to Dr. John G. Read, Associate Professor of Science Education, School of Education, Boston University, for his guidance and inspiration in the preparation of this thesis.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Statement of the problem. The purpose of this thesis is to construct and evaluate an inventory test in general science for the ninth grade using the scientific vocabulary as the basis of the test. Green, Jorgensen, and Gerberich[1] define an inventory test as "a test used as a preliminary check on the degree of mastery existing prior to instruction."

It is hoped that this instrument will serve as an aid in determining the content of preceding science courses; a means of recognizing individual differences in the scientific background of ninth-grade pupils; and an aid in the planning of a more effective course of study.

Justification. The content of science courses at the elementary and junior high school levels is still very indefinite. Courses of study are continuously being revised, and the time allotted varies greatly. Some schools are organized on the 8-4 plan, some on the 6-3-3 plan, and others on the 6-2-4 plan. In the schools where the 8-4 and the 6-2-4 plan exist, pupils entering their first year of high school from a wide variety of grammar schools or two-year junior high schools have extremely varied backgrounds.

1/ Harry A. Green, Albert N. Jorgensen, and J. Raymond Gerberich. Measurement and Evaluation in the Secondary School, Longmans, Green, and Company, New York, 1943, p. 644

Victor H. Noll[1] states that

"Many pupils reach junior high school without any previous instruction in science. At the same time we find pupils with from one to six years of instruction in science. Between the extreme on the one hand of continuous instruction since entrance to the first grade we find all degrees of previous preparation in science. Science instruction in junior high school, then, must either ignore differences in previous preparation . . . . or differentiate the instruction according to previous training in science, which is not easy nor even always possible."

In the Forty-sixth Yearbook[2] we find

"There is better agreement in the total content for the whole junior high school than there is in the grade placement of details for a particular year . . . . Psychologically the great barriers to a satisfactory grade placement of content for science courses are in the large variation in individual abilities and in the varying rates at which different pupils mature. There are also factors of opportunity and the variations in physical and social environment."

Another factor which affects the pupils' science background is the science training of the teacher. It is agreed that the science training of elementary school teachers is often inadequate.

In order to avoid undue repetition of course content and to get a picture of the science experiences of ninth-grade pupils early in the school year, some device is necessary.

1/ Noll, Victor H. The Teaching of Science in the Elementary and Secondary Schools, Longmans, Green and Co., New York, 1939, p. 99

2/ Forty-sixth Yearbook of the Society for the Study of Education, Science Education in the American Schools, University of Chicago Press, Chicago 37, Illinois, p. 157

The Thirty-first Yearbook[1] states that

> "At the outset of a new unit of work the teacher may provide the means of making an inventory of the pupils' present acquisition of the learning products which the unit is planned to give. The inventory is successful to the degree that it discovers the individual background, for only with such individual diagnosis can the teacher proceed intelligently with the work of the unit."

Among the very limited number of tests available in the field of general science, achievement tests predominate. Only one test was found which seemed suitable for inventory purposes. This one consists of a series of tests which could be used for unit pre-tests throughout the year.

Green, Jorgensen, and Gerberich[2] state that

> "Adequate science instruction should be expected as a matter of course in an age of science such as the present. It is rather surprising to note, therefore, that progress in the selection and organization of science content and improvement in teaching and testing methods and materials in recent years have been slight in spite of the great practical value of science and its natural appeal to the curiosity and interests of children."

Scope and limitations. The General Science Inventory Test will be given to ninth-grade pupils very early in the school year to help determine their previous science experiences. It is an objective pencil-and-paper test showing relationships between science topics covered in ninth-grade science and vocabulary based on scientific facts and concepts.

1/ Thirty-first Yearbook of the National Society for the Study of Education, Part I. A Program in Teaching Science, Public School Publishing Company, Bloomingdale, Illinois, 1932, p. 213

2/ Op. cit., p. 408

Green, Jorgenson, and Gerberich[1] state that

> "Facts in science are the vehicles for thought. The understanding of relationship of facts and of generalized ideas is deemed most important. . . . Tests should, therefore, so far as possible measure the relational aspects of science."

This test measure only two of the general objectives of science education as listed in the Forty-sixth Yearbook.[2] These are the functional understanding of facts and the development of functional concepts. It covers forty topics representing the major areas of general science study. The vocabulary words included in the test are typical and representative of the science vocabulary used in junior high school science books. The words in the list were checked against the word book of Thorndike and Lorge,[3] the science vocabulary list of Curtis,[4] and words included in nine ninth-grade science textbooks.

The test score will give a comparative picture of the extent of the pupils' general science background, and reference to the individual sections would be valuable throughout the year in planning units to which the topics relate. It will

1/ Op. cit., p. 414

2/ Forty-sixth Yearbook of the National Society for the Study of Education, Part I. Science Education in American Schools, University of Chicago Press, Chicago 37, Illinois, p. 142

3/ Edward Thorndike and Irving Lorge. The Teachers Word Book of 30,000 Words, Bureau of Publications, Teachers College, Columbia University, New York, 1944

4/ Francis Curtis. Investigation of Vocabulary in Textbooks of Science for Secondary Schools, Ginn and Company, Boston, 1938 also be valuable in locating common misconceptions which would definitely hinder normal progress. The test could also be used at the end of the eighth or ninth grade to determine pupil achievement, although this is not the original purpose of the test.

The test is not a simple matching test of memorized facts. Relationships must be made which require elements of the inductive method. The relationships made by ninth-grade pupils will include their science experiences previous to taking the test. Whitehead[1] makes this clear when he writes

> "It is impossible to overemphasize the point that the key to the process of induction, as used either in science or in our ordinary life, is to be found in the right understanding of the immediate occasion of knowledge in its full concreteness. . . . . Inductive reasoning proceeds from the particular occasion to the particular community of occasions, and from the particular community to relations between particular occasions within that community."

The Forty-fifth Yearbook says of facts and understandings

> "We may say that a pupil understands when he is able to act, feel or think intelligently with respect to a situation.[2] . . . . The pupil must develop worthwhile understandings of the world in which we live as well as the symbols associated with this world.[3] . . . .

1/ A. N. Whitehead. <u>Science and the Modern World</u>, Pelican Mentor Books, New American Library of World Literature, New York, 1925, p. 44

2/ Forty-fifth Yearbook of the National Society for the Study of Education, Part I. <u>The Measurement of Understanding</u>, University of Chicago Press, Chicago, Illinois, 1946, p. 28

3/ Ibid., p. 34

Symbols acquired without being associated with the realities for which they stand are meaningless. As such they are quickly forgotten or, if retained, they are limited in their usefulness to situations which are practically identical with those in which they are learned.1/. . . . To provide evidence of understanding, evaluation situations must contain an element of novelty, but not too much novelty. . . . . Understanding is attested when the pupil dips into his knowledge and fits it into a new pattern of thought or action which could not have been directly learned."2/

The Thirty-first Yearbook3/ states

"It should also be planned that the generalizations become associated as the course proceeds into larger and larger ideas as the different aspects of the environment are considered and as newer relationships may come to light through learning. In such development of the learning outcomes the detailed facts, the terms for materials, forces, processes and phenomena, and the associated ideas will be woven into a systematized, though small group of generalizations, which will be those knowledges most usable in interpreting life situations and in assuring the desirable adjustments sought as the result of the work in science on this level."

In view of the varying science backgrounds of pupils entering ninth grade under the 8-4 and the 6-2-4 plan and the absence of any suitable means of determining the extent of that background, the general science teacher needs some instrument to use at the beginning of the year which will give him a better understanding of his classes and of the individual

1/ Ibid., p. 36

2/ Ibid., p. 46

3/ Op. cit., p. 200

differences of class members. Through the use of such an instrument a more effective course of study can be planned; a more efficient use of time will be possible by omitting needless repetition; and greater pupil interest and learning may result.

Dr. Roy O. Billett[1] states that

> "In the introductory phase of a valid teaching-learning cycle teacher activity is dominated by two purposes: (1) to know each pupil in the class as an individual; and (2) to give each pupil reason to believe that he is to have a chance to engage in worthwhile activities."

__Restatement of problem.__ The problem of this thesis is to construct and evaluate an inventory paper-and-pencil test in general science using the science vocabulary for ninth-grade pupils. It is to serve as a means of determining the extent of scientific experiences possessed by individual pupils at the beginning of their ninth-grade science.

1/ Roy O. Billett. __Fundamentals of Secondary School Teaching with Emphasis on Unit Method__, Houghton-Mifflin Company, Boston, 1940, p. 599

CHAPTER II

PRESENT STATUS OF GENERAL SCIENCE TESTS

With so little agreement in the elementary school regarding the course of study in general science, it is not strange that there is much to be done in the field of evaluation at this and higher levels. When the objectives, scope, pupil experiences and grade placement of science content are so poorly defined, evaluation becomes more difficult with each succeeding grade.

The number of available general science tests for grade nine is very limited. Most of the tests are designed to measure achievement and seem difficult for the many classes with which the writer is acquainted. No test was found which was designed primarily as an inventory test.

The following tests for grade nine were found:

Glenn-Gruenberg Instructional Tests in General Science[1/] are designed for diagnostic and inventory purposes. There are thirty-five tests of from seventeen to seventy-two multiple choice, completion, true-false, and matching items. The tests follow typical general science units and are most suitable as a teaching device.

Powers General Science Test[2/] has two forms and includes one hundred multiple choice items. It is intended for use through eighth grade.

1/ World Book Company, Yonkers-on-Hudson, New York, 1932.

2/ Bureau of Publications, Teachers' College, Columbia University, New York, 1927

Ruch-Popenoe General Science Test[1] has two forms with fifty multiple choice and eighty completion items. It is intended for grades eight and nine. It is definitely an achievement test with items divided as follows: biology, thirty per cent; chemistry, twelve per cent; physics, thirty-eight per cent; earth science, twenty per cent.

Cooperative General Science Tests[2] are in three forms with about one hundred multiple choice items giving four or five choices. They cover the field of general science very well and would be very useful as achievement tests for pupils in the college or scientific course.

Stanford Achievement Test[3] is a multiple choice test for grades four through nine. It emphasizes biological science at an elementary level with only a few items at ninth-grade level.

With so few tests of such a limited nature available, the need for more general science tests is evident. The test to be constructed will include forty general science topics most commonly included in nine general science textbooks and two hundred and sixty-four vocabulary words relating to the topics.

1/ World Book Company, Yonkers-on-Hudson, New York, 1923

2/ Cooperative Test Service, New York, 1937-1940

3/ World Book Company, Yonkers-on-Hudson, New York, 1943

CHAPTER III

CONSTRUCTION AND ADMINISTRATION OF THE TEST

Introduction. There are many factors to consider in the construction of any test. The objectives must be determined, the most suitable test style selected, and items written which are clear, concise, and accurate. The test must be long enough to widely sample the pupil's knowledge, its working time should be brief, and it should be easy to administer and score. Without a carefully constructed test to begin with, the administering of the test and the item analysis have little value.

Purpose of the test. This test is intended to provide general science teachers of grade nine with an instrument which can be used at the beginning of the school year to determine the extent of previous science experiences of the pupils entering their classes. As an inventory test it is based on the content of the ninth-grade science course as covered by science textbooks.

General objectives for science. The major objectives for science have been outlined and discussed in detail by the Forty-sixth Yearbook.[1] Because of the number and variety of objectives listed, this test will include only the functional understanding of facts and concepts.

Specific objectives. An inventory test for ninth-grade science would have for its specific objectives those included in the course of study for that grade. With the content and grade

[1] Op. cit., pp. 19-33, 141-150

placement of science courses still so unsettled, a detailed list of accepted specific objectives is a research problem in itself. The objectives for this test will be given in terms of science areas, topics, and vocabulary based on the content of ninth grade science textbooks as listed in the bibliography.

The nine areas of science education have been listed as follows:

Machines and Energy
Chemical and Physical Changes
Sound, Heat, and Light
Electricity
The Universe
Earth and Its Atmosphere
Living Things
Man's Control of His Environment
Social Aspects of Science

The following topics were selected from the textbooks examined:

air	moon
airplane	photography
automobile	physical change
balloon	planets
chemical change	plants
digestion	power
disease	radio

earth	rocks
electricity	ship
energy	soil
fire	sound
first aid	stars
food	sun
fuel	telegraph
heat	telephone
human body	time
insects	universe
light	water
machines	weather
matter	work

The vocabulary words included in the test were selected from a list of seven hundred and sixty words compiled by the writer from glossaries and indices of nine general science books. All words were omitted that were not found in at least five textbooks and checked against the Teacher's Word Book of Thorndike and Lorge[1] and the science vocabulary list of Curtis[2]. The final selection of the two hundred and sixty-four words in the test will be discussed in more detail later. The words selected are listed below.

1/ Op. cit.
2/ Op. cit.

acid
aeration
aileron
alkali
alloy
altimeter
anemia
anemometer
antenna
antibodies
antitoxin
aqueduct
artery
astronomy
atmosphere

bacteria
ballast
barometer
Bell
breaking
bronchial tubes
buoyancy
Burbank
burning

calendar
calorie
carbohydrate
carbon
carburetor
caterpillar
charcoal
chassis
chemical change
chlorine
chlorophyll
chronometer
circuit
clouds
cocoon
coke
combustion
comet
compound
condensation
conductor
constellation
contagious
convection
corona
current
cyclone
cylinder

degree
delta
dew
dew point
diaphragm
dirigible
dislocation
distributor
Draco
dry cell

echo
eclipse
effort
electromagnet
element
equinox
erosion
evaporation
expansion

feldspar
fermentation
fertilizer
fever
flame
flint
food
foot candle
foot pound
fracture
freezing
frequency
fuel
fulcrum
Fulton
fungus
fuse
fuselage

gall bladder
gear
generator
glider
goiter
gondola
granite
grasshopper
Greenwich

hammer
hemorrhage
"high"
horse
humus
hurricane
hydroelectric
hydrogen
hygrometer

immunity
inclined plane
incubation
inflammable
intestines
iron
irrigation

Jupiter

kilowatt hour

larva
larynx
lava
lens
lever
light
lightning
limestone
liver
loam
lungs

maggot
magnetism
marble
Mars
measles
mechanical
Mercury
mercury
meteor
meteorology
microphone
Milky Way
mixture
moon
Morse
moss
moth
mulch
muscle

nebula
negative
neon
nimbus
nitrogen
nutrition

orbit
Orion
oxidation
oxygen

Pasteur
pasteurization
pendulum
perennial
periscope
photosynthesis
Picard
pitch
planet
planetoid
Pluto
Polaris
pollination

potential
precipitation
pressure
prism
propeller
protein
pulley
pulse
pupa

quarantine
quartz

radiant
red corpuscles
reflection
relative humidity
reservoir
rickets
rudder
rusting

saliva
sand
sandstone
satellite
sediment
sedimentary
serum
Sirius
skeleton
smoke
solid
solution
sounder
souring
spark plug
spectrum
spinal cord
stalactite
standard
star
stimulant
stomach
stratosphere
submarine
sun

tearing
tides
thermometer
tourniquet
transformer
translucent
transmitter
transparent
transpiration
turbine

ultraviolet

vegetation
vein
ventilation
Venus
vibration
vitamin
volt

water clock
wave length
wedge
well
wheel and axle

yeast

15 pounds
30 inches
H_2O
212 degrees
62.4 pounds
1,100 feet per second
186,000 miles per second
75th meridian
29½ days

Construction of the test. In selecting the style of the test, an attempt was made to use a test style which would cover a wide variety of items, require a minimum of the pupil's time, and introduce an element of novelty. The complete test is submitted for your examination at the end of this chapter.

In constructing the test with two hundred and sixty-four items relating to forty specific topics which requires only forty minutes to administer, much time and thought were necessary. Each of the eight sections consists of five numbered science topics which relate to the different areas of science education and have no direct relationship to each other. Words relating to each of these topics were listed, and from these lists words were selected which have no direct relationship to the other four topics. Very careful selection of both topics and words was necessary to avoid overlapping of relationships. Some zero items were added which have no relationship to any of the five topics, giving six choices of answers instead of five. The words to be included in each section were randomly arranged, so that no planned order of answers would appear.

The completed test was then submitted to a jury of four people in the field of science education in an attempt to locate overlapping relationships which had escaped the writer's observation. The improved test was then submitted to a seminar group for further study and discussion. It was constantly necessary to keep in mind the fact that relationships made by those with a college science education would not necessarily be the same as those made by ninth graders with their more limited experiences. To verify this, the test was given to a typical group of ten ninth graders between seminar classes and the results reported and discussed. A few more items were removed, so that no item would have more than one acceptable answer.

<u>Administration of the test</u>. In order to obtain necessary data for item analysis and further test evaluation, the inventory test was duplicated in sufficient quantity to distribute to all who were to take the test. Four schools of varying sizes in different localities and seven experienced science teachers cooperated in giving the test to three hundred and seventy-two ninth-grade pupils in fourteen science classes.

Directions on each test with a sample of ten items made further explanation by the teacher unnecessary. Forty minutes was allowed for the test and with very few exceptions proved sufficient. As it is a power and not a speed test, time should be allowed for completion if possible. The usual technique followed is to allow all pupils who intend to finish time to finish the test.

NAME ______________________ HOME ROOM __________ SECTION __________

GENERAL SCIENCE INVENTORY TEST

DIRECTIONS. Put the number of the word at the top of each section beside all the words below which relate directly to it. If a word listed has no direct relationship to any of the five words, put a 0 beside it. Each word listed will have only one number beside it. Omit any you do not know. STUDY THE SAMPLE CAREFULLY.

SAMPLE.

1. Light 2. Weather 3. Fire 4. Seasons 5. Birds

4 spring
5 robin
3 coal
2 pressure
1 electric

0 Big Dipper
3 oxygen
1 reflected
5 blue jay
0 chalk

SECTION I.

1. Insects 2. Balloon 3. Sun 4. First Aid 5. Electricity

___ current
___ fracture
___ dry cell
___ degree
___ grasshopper
___ conductor
___ hemorrhage
___ delta
___ planet
___ kilowatt hour
___ 15 pounds per sq. in.

___ larva
___ volt
___ mixture
___ erosion
___ fuse
___ transparent
___ generator
___ stratosphere
___ eclipse
___ charcoal
___ tourniquet

___ pupa
___ corona
___ Picard
___ transformer
___ star
___ precipitation
___ caterpillar
___ maggot
___ wheel and axle
___ dew point
___ lightning

SECTION II.

1. Heat 2. Machines 3. Moon 4. Plants 5. Rocks

___ eclipse
___ lens
___ stalactites
___ perennial
___ sedimentary
___ clouds
___ chlorine
___ lava
___ convection
___ fungus
___ lever

___ pollination
___ photosynthesis
___ inclined plane
___ artery
___ chlorophyll
___ goiter
___ planetoid
___ diaphragm
___ cocoon
___ tides
___ hammer

___ granite
___ lungs
___ pulley
___ sandstone
___ thermometer
___ satellite
___ 29½ days
___ feldspar
___ antiseptic
___ prism
___ mercury

SECTION III.

1. Telegraph 2. Soil 3. Chemical change 4. Human Body 5. Stars

___element	___spinal cord	___Morse
___burning	___sun	___freezing
___souring	___rusting	___rudder
___gondola	___muscle	___loam
___sediment	___periscope	___circuit
___constellation	___Polaris	___Orion
___skeleton	___acid	___Sirius
___sounder	___electromagnet	___atmosphere
___Mercury	___wedge	___Milky Way
___spark plug	___evaporation	___red corpuscles
___humus	___erosion	___hurricane
___astronomy	___vein	___bronchial tubes

SECTION IV.

1. Energy 2. Matter 3. Telephone 4. Digestion 5. Weather

___barometer	___cyclone	___relative humidity
___fulcrum	___Burbank	___Bell
___stomach	___"high"	___atmosphere
___liver	___saliva	___radiant
___dislocation	___anemometer	___flame
___carburetor	___mechanical	___intestines
___Pluto	___orbit	___alkali
___magnetism	___solid	___serum
___light	___transmitter	___potential
___pulse	___meteorology	

SECTION V.

1. Water 2. Photography 3. Fuel 4. Universe 5. Automobile

___fermentation	___meteor	___flint
___dew	___calendar	___gall bladder
___astronomy	___212 degrees	___spark plug
___reservoir	___food	___comet
___negative	___cylinder	___coke
___chassis	___62.4 pounds	___nebula
___H_2O	___well	___hygrometer
___buoyancy	___moth	___yeast
___gear	___stimulant	___30 inches
___lens		

SECTION VI.

1. Light 2. Airplane 3. Physical change 4. Food 5. Time

___pasteurization
___quartz
___Greenwich
___calorie
___sediment
___pendulum
___equinox
___foot candle
___propeller
___irrigation
___nutrition
___ultraviolet
___tearing

___carbohydrate
___spectrum
___altimeter
___nimbus
___standard
___Draco
___reflection
___foot pound
___condensation
___fuselage
___month
___expansion
___ballast

___75th meridian
___breaking
___smoke
___glider
___translucent
___iron
___aileron
___acid
___antibodies
___186,000 miles per second
___vitamin

SECTION VII.

1. Radio 2. Power 3. Planets 4. Air 5. Ship

___stratosphere
___turbine
___chronometer
___hydroelectric
___Jupiter
___Fulton
___nitrogen
___microphone
___hydrogen
___pressure
___pendulum

___barometer
___mulch
___aeration
___vegetation
___frequency
___ventilation
___water clock
___wave length
___horse
___moons
___rudder

___oxygen
___limestone
___antenna
___fertilizer
___atmosphere
___marble
___star
___moss
___neon
___Venus
___sand

SECTION VIII.

1. Fire 2. Work 3. Sound 4. Earth 5. Disease

___bacteria
___immunity
___larynx
___solution
___combustion
___compound
___measles
___oxidation
___inflammable
___pitch
___quarantine
___transpiration

___alloy
___contagious
___Pasteur
___fuel
___1100 feet per second
___rickets
___carbon
___fever
___planet
___chemical change
___aqueduct

___effort
___protein
___echo
___antitoxin
___anemia
___incubation
___dirigible
___Mars
___distributor
___foot pound
___submarine
___vibration

CHAPTER IV

ANALYSIS OF DATA

Introduction. The evaluation of this instrument includes some general measures of central tendency and variability, an individual item analysis for difficulty and discrimination indices, correlation with I. Q. and with a standard general science achievement test, and determination of estimated reliability. The test has curricular validity.

Statistics on central tendency and variability. The summarized data obtained from the original administration of the test include the arithmetic mean, standard deviation, range, number of items, and number of pupils for the whole test.

TABLE I

RESULTS OF THE TOTAL TEST SCORES

Mean	Standard Deviation	Range	Number of Items	Number of Pupils
130.515	27.15	33-219 or 186	264	372

Individual item analysis. The item analysis is for the purpose of securing a difficulty index and a discrimination index for each test item. In order to obtain the necessary data for the analysis, the test was given to 372 pupils in four schools. The upper 27 per cent and the lower 27 per cent of the scores were used as suggested by Davis,[1]

A tally was made of both the high and low papers to determine the number who answered each item correctly and then to determine the number who answered each item incorrectly. A tally was also made to determine the number who reached the item in each group.

In order to correct for guessing and to obtain the necessary percentages with which to enter the Item Analysis Chart[2] the following formulas were used:

$$P_H = \frac{R_H - \frac{W_H}{K-1}}{N_H - NR_H}$$

$$P_L = \frac{R_L - \frac{W_L}{K-1}}{N_L - NR_L}$$

Where:

R_H = number of testees in the highest 27 per cent that answered the item correctly

W_H = number of testees in the highest 27 per cent that answered the item incorrectly

R_L = number of testees who answered the item correctly in the lowest 27 per cent

W_L = number of testees who answered the item incorrectly in the lowest 27 per cent

N_H = number of testees in the highest 27 per cent of sample

1/ Frederick B. Davis. Item Analysis Data--Their Computation, Interpretation and Use in Test Construction, Graduate School of Education, Harvard University, Cambridge

2/ Ibid.

NR_H = number who did not reach the item in the highest group

NR_L = number who did not reach the item in the lowest group

K = number of choices

The table for the individual item analysis gives the test item; the per cent of right and wrong answers with the highest group of 27 per cent; the per cent of right and wrong answers with the lowest 27 per cent; the corrected percentage for each group; the discrimination index and the difficulty index for each item.

The discrimination index expresses the degree to which a test item distinguishes between pupils of limited ability and those of greater ability. Davis[1] states that any item with a discrimination index of over 20 is significant, and includes a table to show the relationship between his indices and the critical ratio which is often used. The indices in the chart range from 0 to 100.

The difficulty index is a means of comparing the difficulty of a test item with other items in the same test. The index is expressed in numbers from ~~0~~ 1 to ~~100~~ 99 beginning with the most difficult. Items with indices of 50 are most valuable, while those near 0 and 100 are useless. The usual practice is to include items with a wide range of difficulty.

Any negative indices found in the following tables resulted when the original percentages used to enter the Item Analysis

1/ Ibid.

Chart[1] were corrected for guessing and for failure of pupil to reach the item. They may be interpreted as zero items and are not significant.

1/ Ibid.

Table II

INDIVIDUAL ITEM ANALYSIS

Section I

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
current	95	3	85	8	92	83	11	74
fracture	98	1	92	1	98	92	15	85
dry cell	74	24	38	46	69	29	24.5	49
degree	45	49	24	56	35	13	19	35
grasshopper	97	1	98	2	97	98	-4	91.5
conductor	60	40	39	44	52	30	14	45
hemorrhage	93	7	56	21	92	52	34	62
delta	81	8	41	23	79	36	29	54
planet	91	9	82	11	89	80	9.5	71.5
kilowatt hour	95	4	79	12	94	77	20	72.5
15 pounds	59	38	32	49	51	22	19.5	42.5
larva	49	46	14	58	40	2	44	33
volt	98	2	81	13	98	78	31	75
mixture	78	22	42	24	74	37	24	53
erosion	69	28	24	36	63	17	31	44.5
fuse	96	4	86	5	95	85	12.5	76.5
transparent	44	49	20	54	34	9	22.5	33.5
generator	95	5	66	12	94	64	29	67
stratosphere	28	69	9	69	14	-5	30	20
eclipse	96	4	85	7	95	84	16.5	76.5
charcoal	93	7	43	26	92	38	42	58
tourniquet	67	26	34	29	62	28	22	47
pupa	61	33	21	39	54	13	30	41
corona	14	76	7	51	-1	-3	10.5	-7
Picard	19	70	8	46	5	-1	16.5	10.5
transformer	83	17	53	27	80	48	22	58
star	85	14	69	23	82	64	14	63
precipitation	63	30	30	30	57	24	21.5	45
caterpillar	97	1	83	8	97	81	25.5	76
maggot	68	24	41	28	63	35	18	4.9
wheel and axle	84	16	41	31	79	35	29.5	54
dew point	57	31	28	31	51	23	19	43
lightning	97	3	78	19	96	74	27	72

Table II (Continued)

Section II

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
eclipse	92	7	80	12	91	78	14	71.5
lens	36	61	25	50	23	15	32	5
stalactite	34	55	7	37	23	0	38	25.5
perennial	29	57	4	28	18	-2	34	22
sedimentary	42	45	11	22	43	6	34	35.5
clouds	55	43	23	53	46	12	26	38
chlorine	74	17	22	26	69	17	35	46
lava	54	43	25	48	45	15	22	39
convection	83	13	43	21	80	38	28	55
fungus	66	29	23	28	60	17	30	44
lever	91	5	50	20	90	46	34	60
pollination	47	37	12	23	40	7	30	34.5
photosynthesis	32	52	9	31	22	3	27	25.5
inclined plane	60	27	41	19	55	37	11	48
artery	70	25	20	28	65	14	35.5	44.5
chlorophyll	30	49	3	25	20	-2	36	24
goiter	65	26	18	17	60	15	32	43.5
planetoid	15	72	4	42	-2	-4	7	-10
diaphragm	58	33	13	28	51	7	41	38
cocoon	70	27	27	38	65	19	31	46
tides	74	26	30	35	69	23	30	48
hammer	81	17	53	27	78	48	20	57
granite	99	1	72	7	99	71	43.5	71.5
lungs	82	14	41	29	79	35	30	54
pulley	89	9	47	12	87	45	30.5	58.5
sandstone	95	2	83	9	95	81	19.5	75
thermometer	85	14	56	28	82	50	23	59
satellite	67	29	14	36	61	7	43	41
29½ days	71	28	41	28	65	35	19	50
feldspar	30	48	7	26	20	2	30	24
antiseptic	83	11	19	29	81	13	49	48
prism	44	48	12	34	34	5	32	26
mercury	21	75	11	60	6	-1	19	12

Table II (Continued)

Section III

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
element	7	87	7	68	-10	-7	-5	-21
burning	80	19	57	28	76	51	16.5	57.5
souring	81	16	40	26	78	35	28.5	53.5
gondola	78	15	24	19	75	20	37	48.5
sediment	41	51	9	34	31	2	29.5	30.5
constellation	94	6	44	25	93	39	43	59
skeleton	99	0	93	2	99	93	20.5	87
sounder	74	18	34	21	70	30	25	50
Mercury	42	55	10	67	31	-3	44.5	29.5
spark plug	60	36	27	45	53	18	24.5	42
humus	28	60	5	36	16	-2	32	21
astronomy	97	2	61	19	97	57	42	66
spinal cord	96	2	67	13	96	64	34	68
sun	82	16	60	21	79	56	16	59.5
rusting	95	3	54	20	94	50	38	62
muscle	99	1	88	4	99	87	29	81.5
periscope	37	56	8	40	26	0	41	27
Polaris	62	28	18	27	56	13	31	41.5
acid	0	97	6	69	-19	-8	-13	-26.5
electromagnet	62	35	27	26	55	22	22	44
wedge	77	13	16	22	74	12	45	46
evaporation	18	81	7	67	2	-6	7	5
erosion	77	20	32	35	72	25	30.5	49
vein	93	7	68	12	92	66	24	67
Morse	86	12	36	44	84	27	39.5	53
freezing	40	63	11	67	27	-2	42	27.5
rudder	78	16	21	27	75	16	41	47.5
loam	75	21	36	20	71	31	25	51
circuit	78	19	43	27	74	38	23	53
Orion	79	13	44	20	76	40	23	54
Sirius	77	14	42	21	74	37	24	53
atmosphere	24	72	15	55	10	4	11	19
Milky Way	99	0	81	14	99	76	39	74
red corpuscles	98	2	76	9	98	74	34	73
hurricane	71	24	27	48	66	17	33	45.5
bronchial tubes	96	4	68	19	95	64	31.5	67.5

Table II (Continued)

Section IV

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
barometer	93	4	74	13	92	71	21	68.5
fulcrum	53	24	14	22	48	10	31	38
stomach	100	0	79	11	100	77	38	75
liver	81	18	62	25	78	57	14.5	59.5
dislocation	65	28	15	30	59	9	39	41
carburetor	62	30	18	26	56	13	31	41.5
Pluto	75	20	34	29	71	28	27.5	49.5
magnetism	15	80	18	47	-1	9	-23.5	14
light	60	42	32	44	52	23	19.5	43.5
pulse	49	44	10	54	40	-1	50	33
cyclone	80	17	62	14	77	57	13.5	59
Burbank	60	23	16	21	55	12	31.5	41
"high"	59	37	19	35	52	12	30	40
saliva	70	25	21	33	65	14	35.5	44.5
anemometer	84	10	33	17	82	30	36	53
mechanical	52	43	18	39	43	10	27	36.5
orbit	53	36	14	28	46	8	32	37
solid	82	17	46	22	79	42	25	55.5
transmitter	81	17	59	15	78	56	15	59
meteorology	58	32	17	28	52	11	31.5	40
relative humidity	87	4	42	16	86	39	33.5	56.5
Bell	98	1	84	7	98	83	26.5	77.5
atmosphere	90	10	76	9	88	74	13	68
radiant	57	38	16	34	49	9	33	38
flame	33	60	13	53	21	2	30.5	24.5
intestine	90	7	33	19	89	29	44	55
alkali	20	72	16	24	6	11	-8	-21
serum	50	42	8	45	42	-1	51	33
potential	25	62	4	28	13	-2	29	19

Table II (Continued)

Section V

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
fermentation	62	26	12	26	57	7	41	40
dew	94	5	77	10	93	75	20	71
astronomy	92	6	48	17	91	45	36	60
reservoir	92	6	49	16	91	46	35	60
negative	92	6	66	10	91	64	24	66
chassis	61	33	16	19	56	12	32	41
H_2O	99	1	73	16	99	70	44	71
buoyancy	64	26	16	17	59	13	33	42
gear	100	0	77	5	100	76	39	74
lens	90	10	40	36	88	33	39.5	55.5
meteor	93	6	34	35	92	27	49	55
calendar	25	73	23	36	10	16	-7	26
212 degrees	67	32	41	27	61	36	16	49
food	52	47	34	39	43	26	11.5	41.5
cylinder	96	4	54	23	95	49	41	62
62.4 lbs.	42	50	23	24	32	18	11	36
well	97	2	74	9	97	72	32	71.5
moth	97	2	32	15	97	29	56.5	57
stimulant	41	50	69	30	31	63	-20	48
flint	29	69	34	67	15	21	-5.5	31
gall bladder	85	12	19	27	83	14	49.5	49
spark plug	98	1	83	9	98	81	29	76.5
comet	97	3	51	13	96	48	44	62
coke	78	22	47	37	74	40	22	54
nebula	31	56	4	24	20	-1	36	24
hygrometer	9	79	5	37	-7	-2	-13	-14
yeast	74	21	19	27	70	15	38	46
30 inches	55	41	8	42	47	0	53.5	35

Table II (Continued)

Section VI

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
pasteurization	53	44	32	40	44	25	12.5	41.5
quartz	73	20	6	23	69	1	67	41.5
Greenwich	45	51	11	29	35	5	31	32
calorie	89	9	41	18	87	38	35	56.5
sediment	50	38	11	18	42	8	29	36
pendulum	59	34	7	18	52	3	47.5	38
equinox	9	69	4	17	-5	1	0	1
foot candle	60	33	48	16	53	46	4.5	49.5
propeller	99	1	77	3	99	78	37	75
irrigation	39	54	8	40	28	0	43	28
nutrition	85	11	32	14	83	30	37	53
ultraviolet	71	24	6	21	66	2	58	41
tearing	73	25	25	18	68	22	31	47
carbohydrate	87	10	28	30	85	23	43.5	52
spectrum	21	58	5	15	9	21	-13	28
altimeter	84	13	12	20	81	8	55.5	47
nimbus	62	23	6	22	57	2	53.5	38.5
standard	86	12	34	8	84	34	34	55
Draco	64	22	7	14	60	4	49	40
reflection	89	9	44	19	87	43	32	58
foot pound	50	32	4	32	44	-2	52	34
condensation	55	39	16	16	47	14	24.5	39.5
fuselage	79	16	20	23	76	16	42	48
month	93	7	54	20	92	53	33	62.5
expansion	78	17	27	20	76	24	35	50
ballast	49	53	5	20	42	1	51	33
75th meridian	73	24	39	16	68	38	19	52
breaking	78	18	33	18	74	31	27.5	51.5
smoke	55	39	8	49	47	-2	53.5	35
glider	99	0	60	12	99	62	49	68
translucent	32	52	1	27	22	-4	37	25
iron	22	72	11	50	10	1	25	16
aileron	47	38	9	17	39	6	31.5	34
acid	29	64	16	34	16	10	7	26
antibodies	42	37	7	17	35	4	33	32
186,000 miles	47	51	13	66	37	0	48.5	31.5
vitamin	95	5	75	7	94	79	18.5	73.5

Table II (Continued)

Section VII

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
stratosphere	74	24	43	33	69	40	18.5	52.5
turbine	67	29	20	18	61	17	30.5	44
chronometer	15	68	5	27	1	0	0	1
hydroelectric	87	12	44	12	85	45	28.5	58
Jupiter	93	5	81	8	92	85	9	75.5
Fulton	79	19	31	24	75	30	29	51.5
nitrogen	79	20	48	15	75	48	18	56
microphone	89	8	49	22	87	48	29	59.5
hydrogen	13	87	1	65	-4	-13	15	-21
pressure	69	31	36	44	63	29	21	47.5
pendulum	73	22	7	20	68	3	57	42
barometer	75	24	37	36	70	32	24	51
mulch	75	10	10	13	73	8	49	45
aeration	38	49	8	23	28	3	32	28.5
vegetation	73	24	13	31	68	8	46	44
frequency	47	46	8	23	38	3	39.5	32.5
ventilation	90	12	35	23	88	32	40	55
water clock	27	63	3	60	14	-9	30	20
wave length	70	30	35	33	64	30	22	48
horse	75	22	36	26	71	33	24	51
moons	92	5	64	14	91	67	22	67
rudder	93	7	38	17	92	37	42.5	58
oxygen	97	3	77	6	96	82	21	76
limestone	87	11	16	27	85	12	53	49
antenna	90	10	35	17	88	34	39	56
fertilizer	88	9	22	24	86	18	49	51
atmosphere	93	7	57	10	92	59	28.5	64.5
marble	95	3	24	19	94	22	56	54
star	5	94	4	81	-14	-13	1	-26.5
moss	85	11	11	30	83	5	62	47
neon	22	72	7	28	8	1	22	14
Venus	97	3	80	5	96	95	2.5	86
sand	91	6	27	14	90	26	47	54

Table II (Concluded)

Section VIII

Test Item	Highest 27% R	Highest 27% W	Lowest 27% R	Lowest 27% W	Per Cent High	Per Cent Low	Discrimination Index	Difficulty Index
bacteria	97	1	77	6	99	83	33	78
immunity	45	43	7	24	37	2	42	32
larynx	28	60	5	24	16	0	32	21
solution	55	39	14	27	48	10	31	38
combustion	82	10	49	19	82	50	23	59
compound	61	37	6	25	55	1	57.5	37.5
measles	96	2	78	9	98	84	25	78
oxidation	61	35	31	20	55	30	16	46
inflammable	88	10	63	10	88	68	17	66
pitch	38	56	10	36	28	3	32	28.5
quarantine	93	5	48	12	94	51	37.5	62.5
transpiration	48	31	9	22	43	6	34	35.5
alloy	50	38	8	11	43	7	32	36
contagious	92	4	40	13	93	42	41.5	59.5
Pasteur	53	36	21	26	47	17	21.5	40
fuel	79	18	58	13	77	63	10	61
1100 feet	85	10	39	21	85	40	32	56.5
rickets	78	20	31	26	76	30	30	52
carbon	66	29	31	19	61	31	19	48
fever	93	5	71	10	94	79	18.5	73.5
planet	90	8	61	13	90	67	20.5	66.5
chemical change	48	46	23	33	38	18	15	38
aqueduct	53	32	15	13	48	14	25	40
effort	93	5	64	10	94	71	24.5	69.5
protein	73	24	20	22	69	18	35	46.5
echo	91	7	75	5	91	85	7.5	75
antitoxin	44	42	16	16	37	15	17.5	36
anemia	65	24	39	11	52	45	4.4	49
incubation	17	70	6	25	3	1	10.5	7
dirigible	71	18	9	14	69	7	48	44
Mars	43	58	14	49	32	5	28	32
distributor	45	41	7	31	38	1	49	32
foot pound	15	66	14	37	2	9	-16.5	16
submarine	79	13	26	25	78	26	35	51
vibration	91	4	59	13	93	72	22	69.5

With the paragraph on interpretation preceding the table it is easy to select items with sufficient discrimination power to be included in a refined test. One hundred ninety-nine items have indices of over 20 and are therefore significant. The range of difficulty is from 1 to 91.5, from easiest to most difficult. Two hundred and twenty-two of the items had indices between 26 and 75.

The writer thinks that some items with a discrimination index of below 20 may indicate misconceptions which are significant and have a place in an inventory test.

Coefficient of correlation. In order to further evaluate this test, a random sampling of fifty testees was made. The intelligence quotients for these pupils were obtained. The Cooperative General Science Test, Form Q,1/ was given to the same group. The coefficient of correlation between the original test scores and the intelligence quotients was .42 and between the test and the standard test scores was .56.

Reliability. One method for obtaining an estimate of the coefficient of reliability is the "Footrule" coefficient.2/ The only necessary data are the mean and standard deviation of the test scores and the number of test items.

1/ Op. cit.

2/ Green, Jorgensen, and Gerberich, op. cit., p. 591

The formula follows:

$$r_{tt} = \frac{n}{n-1} \quad \frac{\sigma t^2 - n\bar{p}\bar{q}}{\sigma t^2}$$

where: $\bar{p} = \frac{M_t}{N}$ $\quad \bar{q} = 1.00 - \bar{p}$

M_t = arithmetic mean of the test scores

σt = standard deviation of the test scores

n = number of test items

The estimated coefficient of reliability thus obtained is 0.92. Retesting is the best method to use for a test of this type, and that must await further experimentation.

CHAPTER V

SUMMARY AND CONCLUSIONS

<u>Summary</u>. The purpose of this test is to provide teachers of general science nine with an instrument which can be used at the beginning of the school year to help acquaint them with the science background of individual pupils entering their classes.

After carefully selecting test items and constructing the test, the writer submitted it to a jury, a science seminar group, and a small sampling of ninth-grade pupils in order to remove any confusing or overlapping items.

In order to secure necessary data for evaluation of the test it was given to three hundred seventy-two pupils enrolled in ninth-grade general science from four high schools.

The evaluation consisted in obtaining the arithmetic mean, the standard deviation and the range of the total test scores, an individual item analysis for discrimination and difficulty indices, the coefficient of correlation between the test score and I. Q., the coefficient of correlation between the test score and the score from a published standard test, and an estimated reliability coefficient.

<u>Conclusions</u>.

1. The test has curricular validity for ninth-grade general science.

2. The low correlation with the I. Q. and the standard achievement test indicate that the test does not duplicate these methods of evaluation as to purpose or results.
3. The "Footrule" coefficient shows an estimated reliability of 0.92.
4. The discrimination index for the items shows 199 of the 264 items to be discriminating.
5. The difficulty indices show:

   30 items with indices between 1-25
   118 items with indices between 26-50
   104 items with indices between 51-75
   12 items with indices between 76-100

6. Two hundred and sixty-four items and forty science topics give a wide sampling for one 40-minute test.
7. The test is objective, easily administered, and quickly scored.
8. The test could be used as an achievement test at the end of the eighth grade.

CHAPTER VI

SUGGESTIONS FOR FURTHER STUDY

1. Other forms of this test could be constructed and evaluated.
2. With another form of this test the reliability of both instruments could be determined.
3. This inventory test could be given to a larger sampling of pupils to establish norms.
4. Specific objectives for general science courses should be determined.
5. Other types of general science tests for all grades should be constructed and evaluated.
6. Other objective means of evaluation in general science should be made available for those pupils with reading difficulties.
7. A critical evaluation of the general science tests now available could be made.

SELECTED BIBLIOGRAPHY

Beauchamp, Wilbur L., Mayfield, John C., West, Joe Young, Everyday Problems in Science, Scott Foresman and Company, New York, 1940. 752 p.

Billett, Roy O., Fundamentals of Secondary School Teaching, Houghton Mifflin Company, Boston, 1940. 665 p.

Buros, Oscar Krisen, The Nineteen Thirty-eight Mental Measurement Yearbook, Rutgers University Press, New Brunswick, 1938.

________________, The Nineteen Forty Mental Measurement Yearbook, Mental Measurement Yearbook, Highland Park, New Jersey, 1941.

________________, The Third Mental Measurement Yearbook, Rutgers University Press, New Brunswick, 1949.

Carleton, Robert H., Williams, Harry H., Modern-Life Science, J. B. Lippincott Company, New York, 1942. 650 p.

Craig, Gerald S., Lewis, June E., Going Forward with Science, Ginn and Company, Boston, 1947. 413 p.

Curtis, Francis, Investigation of Vocabulary in Textbooks of Science for Secondary Schools, Ginn and Company, Boston, 1938. 127 p.

Davis, Frederick B., Item Analysis Data--Their Computation, Interpretation, and Use in Test Construction, Graduate School of Education, Harvard University, Cambridge, 42 p. and chart.

Davis, Ira C., Sharpe, Richard W., Science--A Story of Progress and Discovery, Henry Holt and Company, New York, 1940. 495 p.

The Measurement of Understanding, Forty-fifth Yearbook, Part I, National Society for the Study of Education, University of Chicago Press, Chicago, 1946. 338 p.

Fowler, George W., Collister, Morton C., Thurston, Ernest L., Our Surroundings, The Iroquois Publishing Company, Inc., Syracuse, New York, 1943. 757 p.

Green, Harry A., Jorgensen, Albert N., Gerberich, J. Raymond, Measurement and Evaluation in the Secondary School, Longmans, Green and Company, New York, 1943. 670 p.

Guilford, J. P., Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, Inc., New York, 1942. 333 p.

Hunter, George W., Whitman, Walter G., Problems in General Science, American Book Company, Boston, 1944. 688 p.

Lemkin, William, Brooks, William O., Visualized General Science, Oxford Book Company, New York, 1939. 544 p.

Meister, Morris, Living with Science, Charles Scribners Sons, New York, 1940. 791 p.

Newman, Barclay M., Haskelberg, Sebastian, Boal, Esther, Vitalized General Science, College Entrance Book Company, New York, 1947. 380 p.

Noll, Victor H., The Teaching of Science in the Elementary and Secondary School, Longmans, Green and Company, New York, 1939. 238 p.

A Program for Teaching Science, Thirty-first Yearbook, The National Society for the Study of Education, Public School Publishing Company, Bloomington, Illinois, 1932. 370 p.

Progressive Education Association, Commission on Secondary School Curriculum, Science in General Education, D. Appleton-Century Company, New York, 1938. 591 p.

Ross, C. C., Measurement in Today's Schools, Prentice-Hall, Inc. New York, 1947. 551 p.

Science Education in American Schools, Forty-sixth Yearbook, Part I, National Society for the Study of Education, University of Chicago Press, Chicago, 1947. 306 p.

Thorndike, Edward, Lorge, Irving, The Teacher's Word Book of 30,000 Words, Bureau of Publications, Teachers College, Columbia University, New York, 1944. 182 p.

Tobin, Louis, The Construction and Evaluation of an Objective Achievement Test in Organic Chemistry for Pharmacy Students Unpublished Master's Thesis, Boston University School of Education, 1948. 47 p.

Whitehead, A. N., Science and the Modern World, Pelican Mentor Books, New American Library of World Literature, New York, 1925. 212 p.