

READABILITY OF GENERAL PSYCHOLOGY TEXTBOOK

MATERIAL; A CROSS-VALIDATION STUDY OF
THE DALE-CHALL READABILITY FORMULA

By

WAYNE DALE LEE

Bachelor of Science in Education
Southeastern State College
Durant, Oklahoma
1955

Master of Teaching
Southeastern State College
Durant, Oklahoma
1958

Submitted to the Faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF EDUCATION
May, 1963

JAN 8 1961

READABILITY OF GENERAL PSYCHOLOGY TEXTBOOK

MATERIAL: A CROSS-VALIDATION STUDY OF
THE DALE-CHALL READABILITY FORMULA

Thesis Approved:

W Ware Marsden
Thesis Adviser

Bernard R Belden

William W Rambo

Robert Muehlen
Dean of the Graduate School

ACKNOWLEDGMENTS

The writer wishes to express his thanks to Dr. Ware Marsden and Dr. Bernard R. Belden, Co-chairmen of his Advisory Committee, for their guidance and encouragement during all phases of this study. He also wishes to express appreciation to the other member of his committee, Dr. William Rambo, and to Dr. Ernest Dewey, Dr. Neil Kent, and Dr. Edward Vodicka, former members of his committee.

Indebtedness is acknowledged to the staff of the Oklahoma State University Reading Center, in particular to Dr. Darrell D. Ray and Mrs. Carol Snelson, who assisted in obtaining data; to Dr. Robert Scofield, Chairman, Department of Psychology, and his staff, for making their classes available for securing research data.

This study is dedicated to my wife, Virginia, and to my children, David, Mark, and Wayne Sue, who have sacrificed some degree of comfort throughout the past three years.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Need for the Study	2
Purpose of the Study	3
Statement of Problem	3
Experimental Hypothesis	3
Basic Assumptions	3
Definition of Terms	4
Scope of the Study	4
II. A REVIEW OF VALIDATION PROCEDURES USED IN READABILITY MEASUREMENT	5
Quantitative Associational Studies	6
Cross-Validational Studies	20
Comparative Studies	22
Summary	27
III. DEVELOPMENT OF THE INSTRUMENT USED TO TEST COMPREHENSION	30
Measuring Comprehension	30
Selecting Test Items	32
The Pilot Study	34
Item Analysis of Data From Pilot Study	36
Method of Determining Reliability	41
IV. COLLECTION, ANALYSIS, AND RESULTS OF DATA IN THIS EXPERIMENTAL DESIGN	43
Administering the Test of Comprehension	43
Description of the Sample Population	44
Administering the Reading Test	45
Criterion of Adequate Comprehension	47
Criterion of Group Comprehension	48
The Scaled Paragraphs	50
Statistical Analysis	51
Statistical Correlation of Obtained Data	53
Summary	56
V. SUMMARY AND CONCLUSIONS	58
General Summary of the Investigation	58
Summary of Test of Comprehension	58
Summary of Reading Test	59
Summary of Results	60
Concluding Statement	60

Chapter	Page
BIBLIOGRAPHY	62
APPENDIX A. Preliminary Test of Comprehension Administered to Determine Item Analysis	71
APPENDIX B. Completed Test of Comprehension to Obtain Data for the Study	91

LIST OF TABLES

Table	Page
I. Description of Population Used in Pilot Study	35
II. Validity and Difficulty Indices of the Preliminary Test of Comprehension	39
III. Validity and Difficulty Indices of the Completed Test of Comprehension	40
IV. Reliability Estimates of Test of Comprehension	42
V. Student Enrollment by Colleges	45
VI. Three Hundred Ninety-Six Readers Classified by Reading Grade-Equivalents	50
VII. Comprehension Percentages of Reading-Grade Levels of 12 Passages of General Psychology Textbook Material .	53
VIII. Product-Moment Coefficient of Correlation of Data Obtained from Correlating Grade-Equivalent Levels of Students with Readability Index of Reading Paragraphs	54

CHAPTER I

INTRODUCTION

Reading is still a chief means by which persons gain information, skills, and entertainment; and the effectiveness with which books, newspapers, magazines, and pamphlets convey this information remains an important problem. Newspaper and magazine writers and publishers are becoming increasingly aware of communicative deficiencies and are showing concern about developing readable materials. Early attempts to make their publications readable by intuitive and rule-of-thumb methods have given way to an acceptance of the objective methods of measuring readability -- methods that have arisen through an accumulation of research.

Webster's Unabridged Dictionary defines readable as: "legible . . . easy to read because interesting or pleasing, that permits or admits of reading." From these definitions, readability may refer to legibility, interest, or ease of reading, or some combination of these. The concept of readability as comprehensibility has received the most emphasis in research. Usually, when the term "readability" is used today, it is this particular concept which is meant.

For Vogel and Washburne (1928) the idea implicit in the readability index of a text was the average amount of reading ability needed to understand the text. Dale and Chall (1948), in explaining readability, defined it as:

the sum total (including the interaction) of all those elements within a given piece of printed material that affects the success that a group of readers have with it. The success is the extent to which they understand it, read it with an optimum speed, and find it interesting.

Since the turn of the century, more and more attention has been devoted to obtaining valid, reliable, and impartial ratings of textbook difficulty.

Need for the Study

This study is an outgrowth of an attempt to evaluate high school science textbooks for readability difficulty. The Dale-Chall Readability Formula was used to measure the readability index. As work progressed and research was undertaken, it became more apparent that there was reason to question the validity of the Dale-Chall Readability Formula as an adequate instrument for this evaluation.

Review of the literature to ascertain the validity of the Dale-Chall Formula failed to support its use beyond the types of material used in its original validation -- health education materials for adults, foreign affairs magazines, and McCall-Crabbs Standard Test Lessons in Reading. A readability formula should be applied only to books similar to those evaluated in the course of its development; therefore, further research is indicated involving secondary and college level materials to determine how valid the Dale-Chall Readability Formula is when used with printed material other than the types used in the original validation, or subsequent cross-validation studies.

Purpose of the Study

The purpose of this study is to ascertain how valid the Dale-Chall Readability Formula is when used to measure general psychology textbook materials used at Oklahoma State University.

Statement of Problem

This study is designed to investigate the problem: How valid are the predicted grade-level placements of passages taken from the textbook used in general psychology at Oklahoma State University when grade-level placements are correlated with an independent criterion of tested comprehension?

Experimental Hypothesis

This study is designed to test the null hypothesis that no significant difference exists between the coefficient of correlation obtained in the original cross-validation of the Dale-Chall Readability Formula and the coefficient of correlation obtained by correlating the readability index of a comprehension test with reading-grade level of certain students enrolled in general psychology at Oklahoma State University who were able to answer three questions out of four on the test of comprehension.

Basic Assumptions

The following assumptions have been made: (1) the reading achievement level of students enrolled in general psychology at Oklahoma State University can be measured; (2) the reading-achievement level can be measured in terms of a grade level; and (3) a valid

instrument for measuring comprehension of certain passages taken from the textbook used in general psychology at Oklahoma State University can be constructed.

Definition of Terms

Terms used in this study are defined below.

Readability -- ease of understanding or comprehension of the printed text.

Predicted grade-level placement -- grade level of difficulty of printed textbook selections or other printed materials measured by readability formulas.

Textbook used in general psychology -- basic textbook chosen by the Department of Psychology to be used by the students enrolled in general psychology during the Fall Semester, 1961, and Spring Semester, 1962, at Oklahoma State University. The text was: Morgan, Clifford T.

Introduction to Psychology, New York: McGraw-Hill, 1961.

Cross-Validation -- an independent criterion of tested comprehension.

Scope of the Study

This study is limited to investigating the validity of the Dale-Chall Readability Formula when used to measure selections taken from the general psychology textbook used at Oklahoma State University and data secured from students enrolled in ten randomly selected sections of General Psychology. The relative import of validity will be determined by a t test of significance to measure level of confidence.

CHAPTER II

A REVIEW OF VALIDATION PROCEDURES USED IN READABILITY MEASUREMENT

¹
2. The idea underlying readability measurement assumes that readers differ in their ability to read and that printed materials vary in difficulty. Matching the right book to the right reader requires the accurate prediction that a certain book is written at the reading level of a particular reader. The earliest approach to matching was based on judgment. The teacher or librarian made recommendations and selections based on his own experience. Some studies reported that teachers ranked materials in an order that corresponded well with empirically determined difficulty. Even though teachers tended to rank well they still underestimated difficulty and many children were given material that was too hard for them to read and comprehend. These inadequate predictions of difficulty gave rise to the search for objective measures.

Chall (1958) reported that the early search for objective techniques sought to reach three major goals:

. . . the discovery of those factors that validly distinguish easy from hard material, a reliable means for measuring such factors, and an expression of some combination of them in terms of the reading ability essential to comprehension.

The goals set forth were achieved by three kinds of studies: surveys of experts' and readers' opinions, experimental studies of one factor, and quantitative associational studies. The experimental

studies contributed to ways of measuring the factors that distinguish easy material from hard. The quantitative associational studies are the most typical of research in readability. The quantitative associational studies produced the readability formulas. ~~Stop~~

The review of the literature will show objective evidence of the validity of various formulas as the authors sought to establish new and better means of measuring the reading difficulty of printed material. An attempt will be made to study validity indices that have been based on different criteria and have measured a wide range of difficulty of material.

Quantitative Associational Studies

The question asked in quantitative associational studies and survey studies is: What makes some materials easy and some hard? The answers are sought through an analysis of the material.¹ Factors studied are the ones arrived at by experience, the results of surveys of opinion, or intuition. The factors are checked by comparison with a criterion.

The internal factors, that is, vocabulary difficulty, sentence length, prepositional phrases, affixes, etc., are presented in quantitative terms and then related to a criterion. The degree of relationship usually is expressed in terms of a coefficient of correlation. The last step is the combining of the most significant internal factors into a formula, usually a regression equation based on a weighting of the most significant internal elements. A review of quantitative studies that have contributed to validity of readability formulas follows.

Lively and Pressey (1923) published an investigation of an new attempt to evaluate the vocabulary burden of textbooks. Using the Thorndike index, 16 different kinds of materials were analyzed.

The authors suggested:

(1) that the vocabulary difficulty or vocabulary burden of a book, or other piece of reading material, may be evaluated by taking thousand-word samplings of the vocabularies used and examining these samplings with reference to the type of word employed, (2) three methods are suggested for summarizing the facts with regard to such a sampling: (a) range of vocabulary, or number of different words per 1000 words sampled, (b) number of words not occurring in the Thorndike list of 10,000 most common words, and (c) weighted median Thorndike "Word Book" index number.

The study of Lively and Pressey was the beginning of attempts to measure the difficulty level of printed material. This study, although significant, failed to quantify the data in a manner that would yield the kind of results helpful to those choosing books for readers.

A significant study in readability was carried out in the Winnetka, Illinois, Schools. Washburne and Vogel's (1926) goal was to express the difficulty of books through their internal factors. Thirty-six thousand seven hundred and fifty widely scattered children reported on all the books which they had read during the preceding year. The ballots which they filled out were brought together, and the results of their judgments made up a list of 700 books. Each book was reported on by 25 children. This list was graded not according to the actual school grade of the children but according to the grade level of their reading ability. The paragraph-meaning section of the Stanford Achievement Test was used as a measure of silent reading achievement. The criterion comprised 1,000-word samples from 150 books taken from the Winnetka book list. Ten internal factors were studied. Four of these ten factors correlated very highly with the ten factors. These four were:

number of different words in a sample of 1,000 words, number of prepositions per 1,000 words, number of words in 1,000 words not occurring in Thorndike's list of 10,000, and number of simple sentences in 75 sample sentences.

Later, Vogel and Washburne (1928) reported the results of their efforts to use an objective means of measuring reading difficulty. Using a criterion based on an empirical evaluation of difficulty and correlating the findings of their study with the criterion, they reported a validity coefficient of correlation of .845. Five factors were included in the 1928 formula. Washburne and Morphett (1938) presented further evidence of validity of the Winnetka Formula. The formula was revised to include only four internal factors. Internal factors weighted in the regression equation were: (1) grade-placement, (2) number of different words per 1,000 words, (3) number of different uncommon words per 1,000 words outside Thorndike's first 1,500 words, and (4) number of simple sentences per 75 sample sentences. A Coefficient of correlation of .86 was reported for the revised formula. Miller's (1946) study gave indirect evidence of the validity of the Winnetka Formula. Miller, using a test of comprehension taken at quarter intervals of each of the 23 Newbery award books, found that only two were at elementary-grade level. The Winnetka Formula was one of several used to measure readability. Elliott (1941), in designing a study to test objective methods of measuring reading difficulty, found the Washburne-Morphett (Winnetka) Formula to lack reliability to the extent that its validity was impaired.

According to Chall (1958), Washburne and Vogel were the first to study the influence of the structural characteristics of the text.

They also were the first to use a criterion based on an empirical evaluation of difficulty. Heretofore, the significance of factors that differentiated easy from hard material had been established by observation and judgment. Washburne and Vogel were the first to use correlation to related the individual factors with the criterion and a combination of factors with the criterion (multiple correlation). The Winnetka Formula was also the first to predict difficulty by grade levels.

Kebock (1927) analyzed every fifth page of seventh-grade history books. Word-difficulty was determined by using Thorndike's word list. Evidence in this study seems to show little variability of word-difficulty. He was able to ascertain a standard deviation and coefficient of dispersion. Kebock made no claim for validity of his method. No analytic formula was given to measure internal factors of readability.

The decade from 1930 to 1940 produced no fewer than 13 validation studies. Johnson (1930) reported a method of determining difficulty of reading matter. Equated groups of 123 students in each group were given two versions of the same test. Each form of approximately 800 words had the same vocabulary except in the case of 83 words which were monosyllabic in one version and polysyllabic in the other. The result of the experiment showed many of those taking the long-word test made lower scores than those taking the short-word test. Johnson claimed validity of his method by showing consistent increase of polysyllables with difficulty of materials. Although Johnson's study was limited in the number of students in the study, results of his study corroborated the belief that monosyllabic words were less difficult to understand than polysyllabic words.

Patty and Painter (1931) studied all the textbooks used in high schools in the state of Indiana. Their findings showed that textbooks used in the sophomore year of high school were more difficult than were the ones used during the other three years. Their formula used the arithmetic mean of the Thorndike index-number of all words in the sample. No claim for reliability or validity is made by the authors. This study by Patty and Painter illustrated the fact that authors and publishers have not considered the readability of their material when this material has been designed for a particular grade-level.

Dale and Tyler (1934) reported a study concerned with predicting the difficulty of materials for adults. Validity was determined by ascertaining the reading ability of Negro adults with limited reading ability. Reading achievement was measured by administering the Monroe Reading Test. Ten out of 29 factors investigated in reading material were correlated with comprehension scores of adults in the study and yielded a coefficient of .562. Realizing that ten factors would be unwieldy to use in measuring readability, Dale and Tyler selected three factors: (1) number of different technical words in the selection, (2) the number of different hard non-technical words in the selection, and (3) the number of indeterminate clauses. These three factors when combined gave a multiple coefficient of correlation of .511. The coefficient of .511 was nearly as high as when ten factors were used. Their study suggested the probability of using fewer factors in measuring readability.

In Ojeman's (1934) study, an investigation was made of the distribution of reading ability among adults, the factors most closely associated with reading difficulty, and the characteristics of materials

at the various levels of difficulty in terms of these factors. To study the distribution of reading ability, a comprehension test was constructed from representative parent education materials and administered to 209 parents attending study groups. An abridged form of the test was administered to an additional group of 156 parents. The relative difficulty of 16 selections of 500 words each was determined experimentally and the variations of quantitative and qualitative factors with difficulty analyzed. Finally, the characteristics of materials in terms of important factors at the lower levels of difficulty were analyzed to assist writers in preparing simplified materials for parents. Ojeman listed the important findings:

1. The reliability of the original reading test as measured by the correlation of even and odd scores and the application of the Spearman-Brown Formula is .96. The reliability of the abridged reading test, consisting of the first two of the three parts of the original test, is .93.
2. As measured by the technique used in this study, all but three of the sixteen selections taken at random from parent education materials required a greater reading ability than the average of subjects who have not continued their education beyond eighth grade. Nine selections required an ability equal to or above that of the average of subjects who have not advanced beyond high school.
3. Of the fourteen factors studied quantitatively in their relation to reading difficulty, the number of simple sentences, number of prepositions, and all factors of vocabulary difficulty correlated above .60.
4. The materials having a difficulty equal to or below the average reading ability of subjects who had not extended their education beyond the elementary school characterized by (a) low vocabulary difficulty, (b) simplified sentence structure as evidenced either by a relatively large proportion of simple sentences, or by a small proportion of words in dependent clauses, and (c) by a small number of prepositional phrases.

Any claim made by Ojeman for validity was based on a combination of correlation and judgment. Ojeman's study was another attempt to find a means for measuring readability. Chall (1958) in her review of readability, reported that:

the difficulty of each passage of Ojeman's study was established by assigning to it the reading-grade equivalent on a standardized test of the readers who were able to answer correctly at least one-half of the comprehension questions on the passage . . . Ojeman was the first investigator to use adult materials for his criterion and adult subjects to establish the difficulty of the passages. He graded the criterion by means of comprehension questions on the selection. He analyzed the materials for qualitative variables.

Ojeman's study was the first study in readability to set a standard required of readers for adequate comprehension.

McClusky (1934) found that "narrative material in fiction was composed of short simple sentences, while passages in psychology were made up of long complex sentences." This fact was quite striking and corresponded to the levels of difficulty reported by him. Evidence of validity comes from test results only.

Gray and Leary (1935) singled out 64 factors of writing style that could be counted objectively and which were suspected of having an effect on reading difficulty. They tested 44 of these factors with a group of 1,670 adults and found 20 factors had important relationship to reading ease. The factors weighted by Gray and Leary break down into two categories: (1) sentence length and the relationship within sentences, and (2) word load. They arrived at a formula using five variables: (1) number of different hard words, (2) number of first-, second-, and third person pronouns, (3) per cent of different words, (4) average sentence length in words, and (5) number of prepositional phrases. Reading level of the 1,670 subjects participating in the study was obtained by administering the Monroe Silent Reading Test. The multiple correlation between the five variables and the criterion was .644. Elliott (1941) found the Gray-Leary Formula to lack reliability to such an extent that the validity was impaired.

In a study by Stone (1938), a comparison of vocabulary burden in terms of ratio of new words to total words (average repetition), average new words per page, and the per cent of sentences complete in one line were studied statistically. Results of Stone's study claimed relative difficulty can be determined by two factors: new words per total number of words and per cent of sentences completed on one line. Stone made no claim of validity in his study but presented more data exhibiting internal factors that seem to be present when reading difficult material.

Lorge (1939) thought that if readability of a passage could be evaluated adequately, the estimate would have two major values: " . . . placing the book on some scale of comprehensibility [and] indicating to writers of books for specialized population, the nature of the difficulty of their product." In his formula, Lorge used three factors: (1) the number of different hard words, (2) the average sentence length, and (3) the number of prepositional phrases. He thought these the best and most practical predictors of passage difficulty. He correlated his formula with tested comprehension of students tested on McCall-Crabbs Standard Test Lessons in Reading, Books II, III, IV, and V. Each of these passages had been standardized on the basis of the number of questions correctly answered by children in terms of scores on the Thorndike-McCall Reading Scale. Lorge found a coefficient of correlation of .77 between the 376 passages of McCall-Crabbs Standard Test Lessons in Reading and tested comprehension of children. Chall (1958) reported that these reading exercises probably were the best criteria so far devised to test reading level of children. Later Lorge (1944) revised his formula. There was no extensive changes made.

Dale and Chall, using Lorge's intercorrelations for purposes of their study, corrected Lorge's calculations and found the multiple-correlation coefficient of the Lorge and Flesch Formulas to be .66. The work by Lorge has made a large contribution to the development of readability formulas. Flesch and Dale and Chall used the statistical analysis obtained by Lorge in securing their own intercorrelations. His formula has been used often in both research and application.

The measure of difficulty of the expression on the printed page is, according to the "idea analysis" method of Morris and Holverson (1938), a measure of commonness of experiences and the words labelling those experiences. The degree of commonness was estimated from three points of view: (1) frequency of occurrence in the life of the individual, (2) frequency of occurrence in the social environment, and (3) frequency of occurrence at levels of intelligence of sophistication on which the majority of people were found. Non-content words were classified by Morris and Holverson as:

1. Simple conjunctions (but not adverbs such as "consequently" when used as conjunctions).
2. Simple prepositions.
3. Pronouns except when the reference is not clear.
4. Articles, as: 'a', 'an', 'the'.
5. All forms of the verb 'to be' and the auxiliary verbs 'have', 'had', 'do', 'did', 'may', 'might', 'can', 'could', 'must', 'shall', 'will', 'should', 'would'.
6. Simple exclamations, as: 'so', 'ah', 'oh'.
7. Proper names, such as those which have no other function except to name characters in fiction. The names of living people or historical persons are content words when they hold meaning in context.

Later, Lorge (1939), using the McCall-Crabbs Standard Test Lessons in Reading as a criterion, reported that when Books I, II, III, and IV were used, the multiple correlation for the Morris and Holverson study was .7419. Including the additional factor of average sentence length

raised the multiple correlation to .7762; adding the number of prepositional phrases raised the coefficient to .7815; adding the composite number of different hard words raised the multiple correlation to .7821.

Yoakam (1939), using Thorndike's Teacher's Wordbook of 20,000 Words, devised a scale to measure readability of printed material. While Yoakam reported no validity index, subsequent studies supported the use of his formula. Stadtlander (1939) presented a study in which he, using Yoakam's Readability Formula, analyzed four selections of fiction, six of historical material, five dealing with science, one based on aviation, three of a geographic nature, and two adapted from legend. Stadtlander constructed a test of comprehension based on materials chosen for use in the study. The New Stanford Reading Test was administered to each student. She found "a relatively high correlation of reliability between the reading achievement level of the student and the gradation level at which the student was able to answer successfully." Results of testing 2,763 children in grade 4, 5, and 6 confirmed that the Yoakam Formula had a high degree of validity on the materials chosen for this study when used at intermediate-grade level. Using the Yoakam Formula for prediction, Figurel (1942) reported a high correlation between the difficulty level of classics used in Pittsburg, Pennsylvania, Schools at grade level 9A and its rating by the pupils. To test the validity of the Yoakam Formula on religion textbook materials, Latimer (1948) constructed a test of reading comprehension. The test was administered to 750 parochial school children. The test of comprehension consisted of three subtests. The coefficients of correlation were .56, .68, and .83 between the criterion and the reading achievement test. These coefficients of correlation were taken to be

significant in determining the validity of the test of religious materials. Presenting a negative report, Elliott (1941) found the Yoakam Formula to "lack reliability to such an extent that its validity is impaired."

Flesch (1943) used Lorge's three elements and two of his own origin to find the relative difficulty of five popular magazines. These elements were: number of affixed morphemes and number of abstract words. Validity of Flesch's procedure rests on the assumption that the five groups of magazines used in assigning level of difficulty actually represents five levels of increasing comprehension difficulty. Wert (1937), Bushwell (1937), and Taylor (1938) reported that typical readers of the magazines in Flesch's study differ correspondingly in educational level and reading ability. The coefficient of correlation between the number of affixed morphemes and number of abstract words for the total distribution of the 375 passages in the study was found to be .7849. Miller's (1946) study gave indirect evidence of the validity of Flesch's Formula. Miller made a cross-validation study using a test of comprehension. Various formulas were applied to the test material, thus giving further information about several formulas.

Later, Flesch (1948) revised his formula. Four internal factors were studied: (1) average sentence length in words, (2) average word length in syllables, (3) average percentage of "personal words," and (4) average percentage of "personal sentences." Again, he used the McCall-Crabbs Standard Test Lessons in Reading as a criterion. Two multiple-regression formulas were computed: (1) using the first two elements and (2) using the last two elements. Classifying the first two elements as Formula A and the last two elements as Formula B gave

coefficients of correlation of .7047 for Formula A and .4306 for Formula B. Formula A alone with a correlation of .7047 had almost as high a prediction value as the combined earlier (1943) formula whose correlation coefficient was .74.

Gilinsky (1948), in a study of the Flesch Formula, used Thorstone's method of "equal appearing intervals" to rate 75 samples of prose, ranging from pulp fiction to technical articles. Correlations between the readability judgments and Flesch counts on both the original and the revised scales ranged from .61 to .84. Gilinsky concluded that the Flesch Formula was a highly valid index of readability. Allen (1952) found that the Flesch, Dale-Chall, and Lorge Readability Formulas were about equal when predicting the comparative readability of film commentaries. Farr and Jenkins (1949) tabled the values for the Flesch "Human Interest" and "Reading Ease" Formulas. These tables were designed to prevent errors in computation.

The formula produced by Flesch probably has had more wide spread use than any other formula. News media have used it extensively to rate their material and have relied upon the formula to control the level of difficulty of their printed matter.

Dolch (1948) took ten series of readers used in grade one to grade six and analyzed each book for word difficulty and average word-length of each sentence. Results showed a gradual progression upward from grade one through grade six in both areas studied. Dolch assumed basal reading series to be the best graded materials. The Dolch technique was not used extensively.

Dale and Chall (1948) found by taking intercorrelations between four style elements and grade-score of pupils who answered one-half

the questions of McCall-Crabbs Standard Test Lessons in Reading, two elements, number of words not on the Dale List of 3,000 Familiar Words and average sentence length a multiple-correlation coefficient of .70. To support the multiple-correlation index reported in the original study, Dale and Chall (1948) found that on 55 passages of health education materials their two-factor formula predictions correlated .92 with the judgments of readability experts and .90 with the reading grades of children and adults who were able to answer at least three questions out of four on thirty of these passages. On 78 passages of foreign affairs from current-events magazines, government pamphlets, and newspapers, the correlation between the predictions of their formula and judgments of difficulty by expert social studies teachers was .90. With just two factors, Dale and Chall obtained a very high relationship between their formula and the criterion. This formula has been applied extensively in many areas and level of printed media. Authors of comparative studies have listed it among the top three formulas.

Wheeler and Wheeler (1948) tabulated a thousand words from page samples at five or 10 page intervals by using the Winnetka Chart for Determining Grade Placement of Children's Books. The grade placement of the words was assigned by use of the Thorndike Word Book of 20,000 Words. The number and percentage of words at each grade level was determined. No effort was attempted to show reliability and validity.

Several formulas have given objective evidence of a coefficient of validity. The first to use a statistical regression equation to weight internal factors was the Winnetka Formula presented by Washburne and Vogel. This formula used five factors as measures of difficulty. A

later revision by Washburne and Morphett presented further evidence of its validity when used to measure juvenile fiction. Reviews and comparative studies continue to support its use. In 1934, Ojeman reported a study in which he evaluated reading ability among adults. He used the criterion of 50 per cent comprehension as a basis for assigning reading level to students, which set the pattern followed later by Lorge, Flesch, and Dale and Chall. In 1935, Gray and Leary examined 64 factors to ascertain those elements showing the most relationship to reading difficulty. Using a regression equation, these authors found that five factors gave a coefficient nearly as high as when more internal factors were used. Several studies have supported the use of the Gray-Leary Formula, but it has not been used extensively in recent years. In 1939, Lorge used only three internal factors and obtained a coefficient of .77. Lorge was the first researcher to use the McCall-Crabbs Standard Test Lessons in Reading as a criterion. These exercises were used later by Flesch and Dale and Chall. Lorge's three factors, being easier to use than the formulas reported earlier, received wide acceptance by those people concerned with evaluating printed matter for its reading difficulty level; however, the Lorge Formula was not validated on material at upper levels and could not make a positive prediction at the upper levels. Flesch, in 1940, adding two internal factors to Lorge's three factors, and using the intercorrelations obtained from Lorge's use of the McCall-Crabbs Standard Test Lessons in Reading obtained a coefficient of .7849. Later revisions reduced the formula to two parts. Numerous reviews and research papers have supported its use to measure difficulty of printed materials. In 1948, Dale and Chall reported a formula

containing only two factors. Subsequent cross-validation studies reported coefficients of .90 and .92 when used in health education and current foreign affairs materials. These validity studies undertaken at adult level have supported its use in many different types of materials at more advanced levels. Dale and Chall warned against its usage beyond its reported data. These quantitative associational studies all reported objective evidence of validity and have been used by various researchers in evaluating difficulty of reading materials.

Cross-Validation Studies

Few studies have been designed to test the validity of readability formulas through the use of an independent criterion of difficulty. Of those studies designed to test validity, most gave evidence of the agreement of the predicted order of difficulty with the order of difficulty determined by some outside criterion. Very few studies gave indication relative to the validity of the grade-placement predictions. Cross-validation studies reported in the literature are presented here with the criterion used to test for validity of a particular formula.

Flesch (1950) found that concreteness appeared to be somewhat more important for readability than the use of short sentences. He observed that his new formula is a rough measure of the level of abstraction. As such, it may be used as a tool in semantic studies, critical reading, literary appreciation, translation, rating of advertising copy, and propaganda analysis. Certain nouns, adjectives, verbs, participles, pronouns, adverbs, all interjections, and the words yes and no were classified as "definite words". Using McCall-Crabbs Standard Test Lessons in Reading as a criterion gave a coefficient of correlation

with the percentage of these "definite words" of $-.554$. Flesch's obtained coefficient, though not large, indicated that his formula would give a measure of difficulty in studies related to word abstraction.

Spache (1953), using the words outside the Dale List of 769 Words and the average sentence length, obtained a multiple correlation of $.818$ on 224 samples of materials designed to be used in primary grades. The present grading of books at the primary level was accepted by Spache as the criterion. The research by Spache was the only readability formula reported in the literature designed to classify books written at primary level of difficulty.

When Farr, Jenkins, and Paterson (1951) used the number of one-syllable words in 100 words instead of syllables found in each 100 words, a new formula simulating the Flesch combined formulas was obtained. A study of 360 hundred-word samples from employee handbooks showed a "high correlation between Flesch's formula" and the formula derived by Farr, Jenkins, and Paterson. The coefficient of correlation was found to be $-.91$. The authors suggested that their new reading ease index may be substituted safely for Flesch's original formula.

In a study by Bergman (1936) the Washburne-Morphett (Revised Winnetka) placement of 70 supplementary readers and story books was compared with teachers' judgments of the difficulty of the books. The range of placement by the teachers was narrower than that of the formula, and the uncorrected coefficient of correlation between the two measures was $.60$. This study by Bergman confirmed an assumption

held earlier by researchers that judgment of teachers was no entirely dependable when teachers were asked to evaluate difficulty of books.

Miller (1946) attempted to confirm the belief that the lack of popularity among elementary-school children of the books awarded the Newbery Prize is related to the difficulty of these books. Method of obtaining data was to choose four samples out of each book at quarter intervals and construct a test of comprehension on each selection. Each of the students participating in the study had been given the Stanford Reading Test. Miller found that of the 23 books selected by the award committee through 1944 only two rated at elementary level. A comparison was made by a cross-validation study. No coefficient of correlation was given in the study. This study, though not designed to test validity of the Winnetka and Flesch Formulas, gave indirect evidence of their validity if the Miller Comprehension Test can be considered a valid criterion.

Some of the formulas are better predictors than others on a particular type of material. Chall (1958) recognized a weakness in readability formulas when she stated:

. . . it is questionable whether the grade-placement arrived at by the application of any one of these formulas can be used to make a definite statement about the suitability of a particular piece of reading matter for a specific level of reading ability . . . much more research is needed in comparing various readability indexes with external criteria and with one another.

Comparative Studies

Studies designed to use more than one formula on the same material and then to investigate the various grade-placement prediction of the different formulas are termed comparative studies. The following

studies have been reviewed in order to discover the relationships that exist between formulas when used to evaluate the same material.

Russell and Fea (1951), using the grade level obtained on twelve juvenile books rated by 63 children's librarians, found that the mean rating of the six formulas -- namely, Flesch, Dale-Chall, Lewerenz, Lorge, Winnetka, and Yoakam -- approximate closely the librarians' mean ratings of difficulty. The intercorrelations, obtained by rank-difference methods, have the following coefficients; the Dale-Chall correlated .28 with Flesch, .56 with Lewerenz, .54 with Winnetka, and .68 with Yoakam; the Flesch correlated .51 with Lewerenz, .07 with Lorge, .39 with Winnetka, and .36 with Yoakam; the Lewerenz Formula correlated .62 with Lorge, .68 with Winnetka, and .43 with Yoakam; the Winnetka Formula correlated .62 with Yoakam. Russell and Fea reported that "the formulas which tend to give the highest correlation with the rest of the group are the Winnetka and the Dale-Chall." Results from this study indicated that the Dale-Chall and Winnetka Formulas gave comparable grade level predictions on juvenile books.

Walther (1943), using a comparative study of magazines, reported:

There is a high correlation between comparative rankings of different issues of the same magazines. The rankings of the September and the January issues showed a coefficient of correlation of .80; the January and May issues, .91; and the September and May issues, .88.

Walther compared the Winnetka and Gray-Leary Formulas when applied to the same issues and samples of popular magazines and reported "a coefficient of correlation of .73." When popular magazines are compared for readability, the Gray-Leary and Winnetka Formulas have a relative high relationship and measure near the same level of difficulty.

Swanson (1948), using Flesch's original and revised, Dale-Chall, Gray-Leary, and Lorge Readability Formulas for grade-level prediction of two versions of a news story, one containing 173 syllables and the other 131 syllables, gave copies to a randomized population of student-veterans. Within 30 hours journalism students conducted interviews to check readership. Statistical analysis showed that the difference between the medians in the number of paragraphs read by the two samples was significant far beyond the .01 per cent level of confidence. The Gray-Leary Formula gave the lowest level of difficulty rating when the five formulas were compared. No coefficients were given for comparison between the five formulas.

Forbes and Cottle (1953), in developing a technique for finding readability of commonly used standardized tests, applied the Dale-Chall, Flesch, Lorge, Lewerenz, and Yoakam Formulas to each of 27 standardized tests. The mean grade-level score of the five formulas for each test was taken as the criterion of readability for the test. When only vocabulary difficulty of these five formulas was compared to the criterion, the Forbes-Cottle Formula correlated .95 with the Dale-Chall; .83 with the Flesch; .84 with the Lorge; .72 with the Lewerenz; and .90 with the Yoakam. Data from this study suggests that the Dale-Chall Formula has the highest relationship to the criterion, the average of the five formulas.

Smith (1952) did a comparative study using the Dale-Chall, Lorge, and Yoakam Formulas. She compared 27 fourth-grade books used in elementary social studies. A comparison of the books using the three formulas revealed:

- a. The three formulas evaluate very close to the same level.
- b. The difference between the Dale-Chall and Lorge average is .26 of a grade.
- c. The difference between the Yoakam and Dale-Chall average is .04 of a grade.

A summary of the study by Smith reveals that in fourth-grade social studies the Dale-Chall and Lorge Formulas have a grade-level prediction differential of approximately one-fourth grade. The Dale-Chall and Yoakam Formulas rate at approximately the same grade level.

One of the objectives by Dunlap (1954) "compared readability measures with each other and with the results obtained on a direct comprehension check." Twenty-four passages were selected to be evaluated for readability. The 12 readability formulas indexing the selections used in the study by Dunlap were: Beal, Boder, Dale-Chall, Flesch (1943), Flesch (1948), Flesch (1950), Gray-Leary Eight Element Formula, Gray-Leary Five Element Formula, Johnson, Kessler, Lorge (1939), Lorge (1948). Two hundred nine students served as the sample population. Davy Crockett was the book chosen for evaluation. From this book 24 passages were selected for use in the comprehension check. The Lorge (1939) Formula gave a coefficient of correlation of .56 when ranked with the direct comprehension check. This was the highest correlation figure of the 12 formulas used in the study. The Dale-Chall gave a figure of .55; Flesch (1948) gave a correlation of .52; Flesch (1943) Formula gave a coefficient of .51; and Kessler's Formula a correlation figure of .51. Apparently the Lorge (1939), the Kessler, and the Flesch (1948) techniques were equally effective measures when the ranking of samples was considered and when used on this type and level of material. The Flesch (1948), Lorge (1939), Lorge (1948), Dale-Chall, and Kessler techniques correlated highly with each other.

Dunlap concluded that "the Dale-Chall, Flesch (1948), Lorge (1939), and Kessler Readability measures are the most practical and most reliable for the material used in this study." Dunlap suggested that more applications of direct comprehension checks, or other similar devices, to students in many grades, using materials for the different school grades were needed.

Several studies have been designed to discover the relationships that existed when readability formulas were applied to the same printed material. The Dale-Chall and Winnetka Formulas gave approximately the same grade-level prediction when used to evaluate juvenile books. The Gray-Leary and Winnetka Formulas had a high relationship when used to measure popular magazines. The Dale-Chall Formula had the highest relationship to the criterion, the average grade-level prediction of the Dale-Chall, Flesch, Lorge, Lewerenz, and Yoakam Formulas, on 27 standardized tests. On 27 fourth-grade social studies textbooks, the Dale-Chall and Lorge Formulas had a differential of about one-fourth of a grade. The Dale-Chall and Yoakam Formulas measured the social studies books at nearly the same level. On the book Davy Crockett, the Lorge (1939), the Kessler, and Flesch (1948) techniques gave the same grade-level prediction. Dunlap believed the Dale-Chall, Flesch (1948), Lorge (1939), and Kessler Readability measures to be the most practical and reliable for material used in her study. Results from these comparative studies suggest that in some cases one or two formulas were equally effective while in other cases the relationship was not close. In cases where a formula not reporting a coefficient was compared favorable to a formula showing a high coefficient of validity, these comparative studies have given indirect evidence of validity.

Summary

Readability formulas as instruments in evaluating printed materials came into use after Thorndike presented his Teachers Wordbook of 10,000 Words. Early research sought to reach three major goals -- the discovery of factors that validly distinguish easy from hard materials, a reliable means for measuring such factors, and an expression of some combination of them in terms of the reading ability essential to comprehension.

Three kinds of research, surveys of experts' and readers' opinions, experimental studies of one factor, and quantitative associational, were conducted during the early years of research in readability. Factors studied in quantitative associational studies were the ones arrived at by experience, the results of surveys of opinion or intuition. The factors were checked by comparison with a criterion. These internal factors were vocabulary, sentence length, prepositional phrases, affixes, or other internal elements.

Later techniques used to validate formulas were cross-validated studies and comparative studies. Cross-validation studies have been designed to test the validity of readability formulas through the use of an independent criterion of difficulty. Comparative studies have been designed to apply more than one readability formula to the same material and from the coefficients obtained from each formula a comparison of grade-level predictions can be made. From this result, researchers have been able to estimate how closely the formulas predict on the material or types of materials in the study.

The Winnetka, Dale-Tyler, Gray-Leary, Lorge, Flesch, and Dale-Chall Readability Formulas have given the most clear-cut evidence of validity. These formulas used correlational methods. Coefficients obtained by multiple-correlation techniques of the combined factors with the criterion passages were approximately .70. When these formulas are applied to materials similar to that of their criteria, they can be expected to classify the material into broad levels of difficulty. An error of about one grade in the average prediction can be expected.

About half of the original reports of each readability technique gave some empirical evidence of its validity. Others have assumed the validity of their techniques, though some based validity on logical grounds. The evidence on validity from cross-validation studies has been limited. There still is a definite need for validation both of relative difficulty and of grade-placement indices for all of the formulas on a wide variety of materials.

On the basis since 1921 of research in readability, some conclusions can be inferred to guide research in the areas of readability. Several formulas seem to show validity. If no index of validity has been ascertained by the author of a formula, what subsequent research justifies its usages? Different formulas yield different results and grade-placement estimates, even when applied to the same material. The formulas vary in relation to the level of material for which they have been intended. The same formula has not been adequate for measuring the reading level of all materials. The formulas have not measured such factors as experience background, interest and purpose in the reader. The formulas have not measured 'concepts'.

Among many suggested studies needed, Chall (1958) lists the following:

Validation studies on textbook materials.

. . . Since the formulas are used by educational publishers and textbook committees for evaluating textbooks for a particular grade, we need to know how valid the predicted grade-placement indexes of the various formulas are when compared to tested comprehension or to teacher or pupil judgment of difficulty.

. . . cross-validation studies on materials at the upper levels of difficulty.

. . . many of the newer formulas, especially Flesch and Dale-Chall, are being used to estimate difficulty of relative complex reading matter. While the formulas are able to differentiate such materials from fairly easy materials, it is still questionable whether finer discrimination at the upper levels will check with tested comprehension or with expert or reader judgment.

Results reported in research and conclusions reached in reviews reveal the type research needed today in readability investigations. The cross-validation study attempted here has been drawn from needed research in readability that was described by Chall.

CHAPTER III

DEVELOPMENT OF THE INSTRUMENT USED TO TEST COMPREHENSION

It seemed necessary that the test constructed from the selected material should represent as many potential elements of difficulty as possible. This seemed essential in order, first, that the criterion of difficulty obtained for each item might truly represent a combination of many influences; and second, that the final identification of elements of difficulty might be as complete and reliable as the scope of the study would allow. The construction of the Test of Comprehension from representative material of a general psychology textbook has been designed to measure comprehension on this material.

Measuring Comprehension

Comprehension is not a single unitary process but a blending of many processes whose totality represents understanding. From mere perception rises the beginning of understanding, manifested by thoughts, feelings, or impulses aroused by the reading material. From this realm of immediate action, a person arrives ultimately at a level of understanding which implies some degree of intellectual discrimination which requires him to distinguish the thought invited by the words from other thoughts more or less like it (Odgen and Richards, 1927).

Two aspects of comprehension directed the construction of test items used in this study: (1) the ability to form a general impression of the total meaning; and (2) the ability to recall specific components in a selection. No attempt was made to measure emotional reaction, except insofar as that outcome may be positively correlated with ability to comprehend the sense of what is read in the form of a general impression.

In this study, the ability to grasp the essential meaning of a selection in the form of a single impression or generalization was measured by the success with which an individual identified the generalization in a series of statements relating to the selection. Two other methods of measurement were not appropriate to the testing situation of this study. Requiring the reader to state orally the impression gained from reading was not practical for the large number of students involved in this study. Also, requiring a written statement of the general impression gained from reading promised to be an invalid measure of comprehension for persons whose facility in written expression might be limited. Some evidence has been reported (Tyler, 1932) to show that ability to formulate a generalization correlates closely with ability to check the best generalization in a multiple-choice test. Tyler reported coefficients of .79 and .85. The ability to identify the best generalizations in a series of statements was adopted in this study as a valid and expeditious means of measuring a reader's ability to get the general sense of a selection in the form of a single impression.

The multiple-choice test was adopted to measure comprehension as it has been defined here. This test used twelve disconnected passages from a general psychology textbook. One response was required for each test item to measure ability to get the general sense of the paragraph and to measure at the same time ability to recall specific information.

Selecting Test Items

Four criteria were met for selecting paragraphs on the assumption that a passage which met all criteria would meet the conditions of this study and, therefore, be a valid test item. The criteria were:

1. Each item must be independent of the content preceding and following it to insure that responses shall not require further acquaintance with the subject matter from which the item is drawn.
2. Each item must show a completeness indicated by the development of a single unit of thought in order that the reader can get a single impression of the general meaning of the items.
3. Each item must contain as many variants in expression as possible to insure that the identification of parts of difficulty in general psychology textbook materials may be reliable.
4. Each item must be as brief as is possible in order not to discourage readers of limited ability and in order to include a wide sampling of material.

The first step in choosing sample passages was to skim through the entire general psychology textbook. All passages that seemed to meet the necessary requirements of this study were marked. After ascertaining whether the passages met the four criteria mentioned earlier, the selections were evaluated by the Dale-Chall Readability

Formula to determine level of difficulty. The number of words in the passages for the test varied from 215 to 360 words.

In the original validation of the Dale-Chall Formula, the predicted grade-levels for 75 passages on foreign affairs from current-events magazines, government pamphlets, and newspapers ranged in difficulty from intermediate-grade through college level. The 55 passages of health-education text ranged in difficulty from fourth-grade through college level (Chall, 1958). In tests of comprehension designed to test persons of relatively homogeneous ability, provisions should be made to provide test items taken from material that has a predicted grade-level placement above what the expected level of the majority of the group will be. Provisions should be made to provide easier materials for those persons who will not measure up to the level of the majority in the group (Garrett, 1959). Passages that appeared to be undesirable for use in this study were eliminated until there remained 12 selections. The desirability of passages rested primarily on suitability of content and the predicted grade-level of difficulty. The predicted grade-level of difficulty ranged from grade-level 10.2 to grade-level 15.6.

The next step in construction of the test of comprehension was to arrange the 12 passages in rank order of difficulty, easiest to hardest. Following this, there was the necessity of preparing questions of multiple-choice type for each of the twelve selections. As a criterion stated in the testing hypothesis in Chapter I, there was a need to have at least four questions for each selection in the final form of the test of comprehension. Because there probably would be test items that would not measure up to the criterion of item-analysis

validity as explained by Garrett (1959), it was decided to include eight multiple-choice items for each passage on the preliminary test.

Garrett (1959) in discussing content validity defines it when he writes:

Most content tests possess content validity. An arithmetic test or a geography test or a reading test, for example, is valid by definition when it contains a sampling of arithmetic problems, geography questions, and paragraphs to be read.

Questions drawn from psychology textbook passages, by definition, can be assumed to contain content validity or face validity.

To insure that the questions contained content validity, a jury of four faculty members in the Department of Psychology of Oklahoma State University agreed to read the passages and questions to determine what questions or distractors needed to be changed or eliminated. Changes were made to correct questionable items when these four faculty members completed their investigation.

The Pilot Study

The Chairman of the Department of Psychology agreed to allow the three summer sections, 1961, of General Psychology 213 to be used in the initial testing. Characteristics of the initial sample population are shown in Table I.

The preliminary test of comprehension was put together in mimeograph form. Directions to students were placed on the cover sheet. The complete preliminary test of comprehension is included in Appendix A.

As a means of getting better cooperation from each student, no mention was made that the students were participating in a doctoral study. Since the students were enrolled in general psychology it was thought that the plan mentioned above would procure better cooperation.

TABLE I
DESCRIPTION OF POPULATION USED IN PILOT STUDY

		Sex			College Classification					Enrollment By College							
	Number	Male	Female	Mean Age	Freshmen	Sophomore	Junior	Senior	Special or Graduate	Agriculture	Arts & Science	Business	Education	Engineering	Home Economics	Graduate	No Previous Psychology
Number	77	35	42	19.36	16	22	19	16	4	2	32	11	11	6	14	1	70
Percent- age		45.45	54.55		20.78	28.57	24.70	20.78	5.17	2.62	41.56	14.28	14.28	7.80	18.18	1.28	91

* Age range extends from age 17 to age 50.

To eliminate any feeling of threat the students might have, they were told that the Department of Psychology was cooperating with this study.

Inasmuch as the students would not be able to read all 12 passages and answer the 96 questions in the allotted 50 minutes, the test was divided into three parts to insure that an optimum number of responses be included for each test item. Test section one began the test with Passage One and worked as far as each person could in the 50 minutes. Test section two began the test with Passage Five and worked as far as each person could, coming back to Passage One after finishing Passage Twelve. Test section three started the test with Passage Nine, coming back to Passage One after finishing Passage Twelve.

Item Analysis of Data from Pilot Study

Probably the most important single measure of the value of a test is the degree of validity possessed by it. Since the validity of the facts procured depends upon the validity of the measure used to obtain them, it is apparently true that if a test does not measure what it purports to measure, the facts obtained through its use are worthless. It is extremely difficult to prove conclusively by direct methods that a test does or does not measure what it is intended to measure.

In the selection of test items for a test intended to measure particular ability, one of the primary concerns of the test-maker is to procure individual items that are valid or discriminating. The discriminating power of a single item refers to the degree to which success on that item by itself indicates possession of the ability which is being measured. An item may be said to be perfect in

discriminating power when every individual who scores successfully on the item ranks higher on a scale of ability than any individual who fails the item. An item is said to have zero discriminating power when there is no systematic difference between the ability of the individuals who succeed on the item and those who fail. Items of all degrees of discrimination may be found between extremes of perfect and zero discriminating power (Lindquist and Cook, 1955).

Many suggestions are given in the literature of test construction relative to procedures for measuring the "merit" of a single item. They range all the way from bi-serial r , which shows the relationship between success or non-success on the item and the criterion measured, to makeshift devices which may approximate an accurate measure of "goodness". Although bi-serial r gives the best index of the discriminating power of test items, since it is based upon all the data from a group, it has the disadvantage of complexity of calculation not found in other methods (Garrett, 1959).

In the present study the procedure used was discussed by Gray and Leary (1935) and before that was used by Traxler (1932) after being recommended by Holzinger (1928). Garrett (1959) describes the method as "A Short Method of Item Analysis." If the test papers of a group are divided into three classes comprising the highest 25 per cent, the middle 50 per cent, and the lowest 25 per cent of the total scores, then the students in the upper 25 per cent are relatively superior and those in the lower 25 per cent are relatively inferior readers (provided that the entire test is valid and reasonably homogeneous). Using the procedure outlined in Garrett, R_H = number right in high group and R_L = number right in the low group, the discriminative power of an item is simply

$(R_H - R_L)$ or $(R_H - R_L) / N_H$ when written as a validity index. Using the same nomenclature, we may write the difficulty index of an item as $(R_H + R_L) / (N_H + N_L)$ in which N_H and N_L are the numbers in the high and low groups, respectively.

The items having the highest validity indices should be selected for the final test. The four items from each passage having the highest validity index were included in the final test. Since the difficulty index was also determined, the four items included were in ascending order of difficulty. Each item included on the final form of the test of comprehension with its validity index and difficulty index is listed in Table II. Items one through eight have been taken from Passage One. Items nine through 16 have been taken from Passage Two. The other items appearing on Table II have been arranged in subsequent intervals as they were taken from the other passages in the test of comprehension. The items that appear on the final form are noted with the asterisk. According to Garrett (1959), the items to be included on the final test must have a validity index of .20 or more if they are to be considered valid and if the test is fairly long. Table III lists the validity index of each of the 48 items included on the final test. Items one through four have been taken from Passage One. Items five through eight have been taken from Passage Two. The other items have been taken in subsequent intervals from the other passages in the test of comprehension. Validity index and difficulty index are listed in appropriate columns. An investigation of Table III reveals that the validity index has a range from .21 to .87.

TABLE II

VALIDITY AND DIFFICULTY INDICES OF THE PRELIMINARY TEST OF COMPREHENSION

Test Item	Validity Index	Diffi-culty Index	Test Item	Validity Index	Diffi-culty Index	Test Item	Validity Index	Diffi-culty Index	Test Item	Validity Index	Diffi-culty Index
1	.05	.68	25	.56	.82	49	.44	.62	73	.21	.97*
2	.37	.97*	26	.78	.73*	50	.44	.92	74	.16	.70
3	.42	.94*	27	.78	1.00*	51	.58	.84*	75	.16	.89
4	.00	.19	28	.83	.95*	52	.33	.38	76	.00	.00
5	.58	.84*	29	.50	.48	53	.56	.92*	77	.21	.97*
6	.37	.59*	30	.78	1.00	54	.61	.73*	78	.26	.94*
7	-.05	.00	31	.78	.82*	55	.18	.68	79	.37	.89*
8	.26	.97	32	.67	.82	56	.50	.83*	80	.16	.89
9	.21	.69	33	.53	.93	57	.56	.83	81	.32	.91
10	.42	.94*	34	.53	1.00	58	.25	.83	82	.21	.82
11	.26	.78	35	.58	.89	59	.69	.72*	83	.37	.85*
12	.47	.91*	36	.11	.38	60	.63	.76*	84	.32	.94
13	.37	.91*	37	.58	.82*	61	.60	.75*	85	.42	.65*
14	.16	.72	38	.58	.96*	62	.60	.75	86	.32	.82
15	.32	.93	39	.63	.81*	63	.87	.54*	87	.42	.91*
16	.53	.87*	40	.74	.86*	64	.60	.75	88	.53	.90
17	.53	.83	41	.68	.93*	65	.26	.92*	89	.00	.26
18	.58	.93*	42	.63	.96	66	.16	.94	90	.57	.81*
19	.58	.93*	43	.63	.96*	67	.11	.86	91	.33	.60*
20	.56	.96	44	.58	.89	68	.26	.94	92	-.05	.30
21	.74	.96*	45	.53	.93	69	.32	.34*	93	.41	.96*
22	.58	.96	46	.63	.89*	70	.26	.71*	94	.21	.88
23	.72	1.00*	47	.84	.77*	71	.26	.89*	95	.15	.96
24	.50	.83	48	.58	1.00	72	.21	.91	96	.38	.52*

* Test items used in final form of test of comprehension

TABLE III

VALIDITY AND DIFFICULTY INDICES OF THE COMPLETED TEST OF COMPREHENSION

Test Item	Validity Index	Diffi-culty Index	Test Item	Validity Index	Diffi-culty Index	Test Item	Validity Index	Diffi-culty Index	Test Item	Validity Index	Diffi-culty Index
1	.37	.59	13	.78	.73	25	.61	.73	37	.37	.89
2	.58	.84	14	.78	.82	26	.50	.83	38	.26	.94
3	.42	.94	15	.83	.95	27	.58	.85	39	.21	.97
4	.37	.97	16	.78	1.00	28	.56	.92	40	.21	.97
5	.53	.87	17	.63	.81	29	.87	.54	41	.42	.65
6	.37	.91	18	.58	.82	30	.69	.72	42	.37	.85
7	.47	.91	19	.74	.86	31	.60	.75	43	.53	.90
8	.42	.94	20	.58	.96	32	.63	.76	44	.42	.91
9	.58	.93	21	.84	.77	33	.32	.34	45	.38	.52
10	.58	.93	22	.63	.89	34	.26	.71	46	.33	.60
11	.74	.96	23	.68	.93	35	.26	.89	47	.57	.81
12	.72	1.00	24	.63	.96	36	.26	.92	48	.41	.96

Method of Determining Reliability

Certain limitations were imposed because of the nature of the study. Permission was granted to test the students in 10 sections of General Psychology. Two class periods were made available for each section. Because of the necessity to use one period for obtaining reading achievement level and the other period for testing for comprehension of content of the general psychology passages, the only method feasible with these limitations was the split-half technique for determining reliability.

Garrett (1947) confirms that the split-half technique for determining the coefficient of reliability, in this case internal consistency, is thought by many to be the best method. Its main advantage is that the data are collected upon one occasion; hence variations introduced by differences between the two testing situations are eliminated. By deliberate intent a test could be divided into split-halves in such a manner that a higher coefficient of correlation index could be obtained, thus causing a test situation not having a unique value. Garrett (1958) defends the use of the split-half technique by asserting that the previous criticism is true only when items are of equal difficulty. When items are in strict order of merit from least to most difficult, the split into odds and evens give a unique determination of the reliability coefficient (Garrett, 1947).

Wert (1954), in comparing the split-half technique and the subsequent use of the Spearman-Brown Formula to the method of rational equivalence states: "To the extent that the items of the test measure the same function these two methods will yield similar estimates of the reliability of a test."

In the split-half method the test is first divided into two equivalent "halves" and the correlation found for these half-tests. From the reliability of the half-test, the self-correction of the whole test is then estimated by the Spearman-Brown Prophecy Formula. The procedure, in detail, is to make up two sets of scores by combining alternate items in the test. The first set of scores represents performance on the even-numbered items; and the second set of scores, performance on the odd-numbered items. Other ways of making up two half-tests are employed, but the odds-even split is the one most commonly used (Garrett, 1958).

In this study the odds and evens for each student in the sampling population were tallied. The sampling population used to compute the coefficient of internal consistency was the 237 students who were tested in the Fall Semester, 1961. This is a separate population from which the validity index scores were obtained. Spurious results occur when the same populations are used. The coefficient of correlation was calculated from raw or obtained data and a coefficient of .8409 was obtained. Table IV illustrates the n , r before application of the Spearman-Brown Formula, r after use of Spearman-Brown Formula, mean score of odd and even items, and standard deviation of odd and even items.

TABLE IV

RELIABILITY ESTIMATES OF TEST OF COMPREHENSION

	r^2	r	Mean		S. D.		n
			Odd	Even	Odd	Even	
Total	.7256	.8409	18.49	20.25	2.96	2.97	237

r^2 = Pierson r between odd and even samples.

r = r corrected by Spearman-Brown Prophecy Formula.

CHAPTER IV
COLLECTION, ANALYSIS, AND RESULTS OF
DATA IN THIS EXPERIMENTAL DESIGN

When research first is conceived, the researcher must consider what his end result may be. He will need to hypothesize tentative means of securing an answer to his problem as well as tentative results. In any experimental design, certain ground rules must prevail. The person doing the research cannot veer from recognized statistical procedure. To do so adds error to the design to such an extent that little confidence can be placed in the results of the experiment. Chapter IV is designed to report the statistical procedure used in this investigation and to support this procedure by statistical concepts and precedent.

Administering the Test of Comprehension

The test of comprehension contained twelve passages. Each passage required that an attempt be made to answer four questions after the student had completed reading the passage. This researcher had access to ten section of General Psychology at Oklahoma State University. Two fifty-minute sessions were made available for purposes of this study.

Results obtained during the pilot study, used in preparing the Test of Comprehension, revealed that there was a likelihood that some students might not be able to complete the test during the fifty-minute session available for this one test. The Test of Comprehension was designed

to measure power rather than speed. A record was kept of the students unable to answer any of the questions asked for Passage Twelve. There were 37 students who did not respond to any of the questions of Passage Twelve. This in no way affected the results of this study. Passage Twelve was included in the study as being within the comprehension level of people who read at grade 15.

No claim to representativeness of the sampling population can be made when the population is drawn from a select group, in this case, students enrolled in general psychology. Ten sections of general psychology were selected randomly for purposes of investigation. Five were selected from the 17 sections available in the Fall Semester, 1961. Five were selected from the 18 sections available in the Spring Semester, 1962.

Description of the Sample Population

Students enrolled in general psychology and participating in this study represented six of the colleges at Oklahoma State University. Table V gives the per cent of the sample enrolled in each of these six colleges and a breakdown of these colleges by their percentage of total enrollment. There were 216 freshmen, 123 sophomores, 41 juniors, 13 seniors, 2 graduates, and 1 special student participating in this research. One hundred eighty male students and 216 female students cooperated in this study. A range in age from 17 to 51 years was obtained. The age recorded was the age closest to each student's birthday. The mean age was 19.6 years.

There were 510 students who were tested by at least one of the two tests used in this study. Due to changing of schedules, dropping the

class, or being absent on one of the testing days, complete records were obtained for 396 students. An effort was made to collect the testing data exactly fourteen days apart so that the Test of Comprehension was administered fourteen days before the reading test. Only in one section of general psychology was the time separating the testing days as much as 21 days.

TABLE V
STUDENT ENROLLMENT BY COLLEGES

	Agriculture	Arts and Science	Business	Education	Engineering	Home Economics	
College Population By Percentile	11.5	26.6	17.2	11.7	24.1	8.8	
Sample Population By Percentile	2.0	34.3	31.6	17.9	3.2	10.8	

Although no claim is made to representativeness of the total population, a study of Table V will reveal how closely the sample population represents the total population of Oklahoma State University.

Administering the Reading Test

Any test to measure reading achievement must meet the usual requirements of test construction and test standardization. The test should be expressed in grade-levels to effect better understanding in classroom usage. The Nelson-Denny Reading Test, 1960 Revision, was

chosen as a test to measure and establish a reading achievement level of the 396 students included in this study. Directions for administering the test call for 30 minutes of actual working time. There was no problem in administering this test during the 50 minutes of a class session.

The 1939 form of The Nelson-Denny Reading Test gave coefficients of correlation between The Nelson-Denny and Shank Reading Test II, Shank Reading Test III, Traxler Silent Reading Test, and Iowa Advanced Silent Reading Tests that do not differ beyond a probable error of plus or minus .027. All the correlations were above .80 and one was above .90 (Traxler, 1940). The coefficient of reliability for The Nelson-Denny Reading Test, 1939 edition, was .787. The probable error for this same test was plus or minus .024. The index for The Nelson-Denny Reading Test compares favorably with the coefficient of .800 on the Iowa Advanced Silent Reading Test. Implications found by Traxler (1940) suggest that the "results of a single reading test have rather high value for the prediction of the readability of the pupils over a period of at least one year." Held (1939) reported that the University of Pittsburg used The Nelson-Denny Reading Test as a placement test for students enrolled in freshman English. The 1960 Revision of The Nelson-Denny Reading Test did not receive any reviews in Buros' Fifth Mental Measurement Yearbook (1959); therefore, no current research data are available.

In the manual of directions for The Nelson-Denny Reading Test, 1960 Revision, the publisher claims that the internal consistency of the 1939 edition was improved. The equivalent-forms method was used to derive a reliability coefficient. Both A and B forms of the test were administered to the same students. The coefficient of reliability

for the total test was .92. When The Nelson-Denny Reading Test, 1960 Revision, was correlated with the American Council on Education Test, a coefficient of correlation of .830 was ascertained. A correlation between the Co-operative English Test and the 1960 Revision of The Nelson-Denny Reading Test gave a coefficient of .730. The Otis I.Q. and The Nelson-Denny Reading Test, 1960 Revision, coefficient was .551. These criteria support the use of The Nelson-Denny Reading Test, 1960 Revision, as an instrument to measure reading achievement at the high school and college level. Chall (1947) used the 1939 form of The Nelson-Denny Reading Test in measuring the reading achievement of students who took part in her investigation of readability of health-education materials and foreign affairs materials. The comprehension levels of these students answering questions taken from health-education materials were used as a criterion in the original cross-validation of the Dale-Chall Readability Formula (1948).

Criterion of Adequate Comprehension

What should be considered adequate comprehension for each individual in a test of comprehension? To determine the answer to this question, Chall (1947) used the following criterion:

. . . an approximation of what is considered adequate in a normal reading situation would be the best criterion. In a normal situation, we should not expect a reader to be one hundred per cent correct in his reaction to a piece of printed material. His failure to recall one particular item would not necessarily mean that he did not adequately understand what he read.

If we considered only a perfect score on a paragraph as adequate comprehension a type of reading situation would be set up which does not actually exist in the natural process of reading. We therefore decided to consider a reader's comprehension of a paragraph adequate if he got at least three out of four of the questions correct.

Lorge (1944) speaks of the readability index as signifying the reading grade "at which the average school child will be able to answer . . . questions on a comprehension test." Flesch (1943) writes: "the estimate is expressed in a figure that indicates the reading-grade at which the average school child will be able to answer." In using the comprehension level of 75 per cent, Flesch (1943) writes:

The statistical readability formula is a means of gauging the ease and interest with which a book, article, or story will be read. The estimate is expressed in a figure that indicates the reading-grade at which the average school child will be able to answer about three-quarters of the questions in a reading comprehension test concerning detail, appreciations, import, vocabulary, and concepts of the text, with adequate completeness and correctness.

In support of this criterion of 75 per cent comprehension, the literature reports that several of the studies designed to use a test of comprehension have used a criterion of 75 per cent. Lorge and Flesch used a criterion of 75 per cent comprehension in the development of their readability formulas. The Dale-Chall Formula, an outgrowth of Chall's (1947) study, used the criterion of 75 per cent. A criterion of 75 per cent comprehension seems to assure that the reader could not get by on his ability to recall details alone. The reader would be faced with one of the difficult types of reading and have to respond correctly.

Criterion of Group Comprehension

The natural reading situation here must be considered. This experiment was designed to ascertain which paragraph would be understood adequately by given levels of reading ability. Should an author write for a specific reading ability group, he would trust that all or nearly all individuals could understand what they read. Chall (1947)

"decided to consider a paragraph within the comprehension of a particular reading group if at least 75 per cent . . . within that group could understand the paragraph . . ." To report that a paragraph is at a level within the comprehension of readers who have ninth grade reading ability means that at least 75 per cent of the pupils with reading grade equivalents ranging from 9.0 to 9.9 were able to answer at least three out of four items correct on that paragraph. This assumes the same paragraph will also be within the comprehension level of students who have reading achievement levels of above grade nine. It was decided to label a paragraph at the lowest grade level in which at least 75 per cent of the individuals could understand the paragraph.

To further illustrate, for Paragraph Eleven:

39 per cent of the students with reading grades ranging from 9.0 - 9.9 were able to understand the paragraph (answered three or four responses correctly).

63 per cent of the students with reading grades ranging from 10.0 - 10.9 were able to understand the paragraph.

68 per cent of the students with reading grades ranging from 11.0 - 11.9 were able to understand the paragraph.

78 per cent of the students with reading grades ranging from 12.0 - 12.9 were able to understand the paragraph.

86 per cent of the students with reading grades ranging from 13.0 - 13.9 were able to understand the paragraph.

86 per cent of the students with reading grades ranging from 14.0 - 14.9 were able to understand the paragraph.

97 per cent of the students with reading grades ranging from 15.0 - 15.9 were able to understand the paragraph.

The criteria of individual and group comprehension described above demonstrates that Paragraph Eleven falls within the comprehension of groups with reading abilities ranging from 12.0 - 12.9, 13.0 - 13.9, 14.0 - 14.9, 15.0 - 15.9, and 16.0 and higher. It is assumed that

Paragraph Eleven would also be within the comprehension level of students who have reading levels above grade 16. A paragraph in which 75 per cent of the students were able to understand it would label the paragraph as the grade level of the students in that group. Any grade equivalent with scores equal to 75 per cent or above would also be able to understand the paragraph.

The Scaled Paragraphs

The criterion of group comprehension was ascertained on the basis of the reading levels of the students who participated in this study. These reading levels had been determined from the results obtained by administering The Nelson-Denny Reading Test, 1960 Revision. The data from the total group was broken up into grade equivalents. Table VI exhibits the reading-grade equivalents and the number of students in each classification.

TABLE VI
THREE HUNDRED NINETY-SIX READERS CLASSIFIED
BY READING GRADE EQUIVALENTS

Reading Grade Equivalents	Number
7.0-7.9	1
8.0-8.9	2
9.0-9.9	13
10.0-10.9	16
11.0-11.9	25
12.0-12.9	64
13.0-13.9	96
14.0-14.9*	86
15.0-15.9*	59
16.0 and higher*	34
*Extrapolated	

Results from the Completed Test of Comprehension of General Psychology textbook materials were tabulated. An International Business Machine card was used for each student. For each paragraph it was possible for each student to record a score of 0, 1, 2, 3, or 4 for each set of questions asked. These data were coded and punched into the card. In addition to the score on the Test of Comprehension, the reading grade score (raw score) for each subtest of The Nelson-Denny Reading Test, 1960 Revision, was coded and punched into I. B. M. cards. With this information coded into cards, it became possible to make use of Data Processing Equipment.

Statistical Analysis

The 396 cards first were separated by the I. B. M. sorter to obtain the reading grade equivalents. Table VI gives the grade equivalents obtained and the number of students in each classification. After the reading grade equivalent groups were obtained, each group was sorted to find the raw test score for each paragraph. To obtain the comprehension of Paragraph One by grade-equivalent 13 (13.0-13.9), all the individuals in that grade-equivalent group were counted. The total number of students who got a total of three and four questions answered correctly was determined. This number was divided by the total number of pupils in that group which was tested on Passage One. If the percentage obtained was 75 per cent or higher, Passage One was considered within the comprehension of reading grade 13.

To illustrate, 59 students with grade equivalents 15.0 - 15.9 were tested on Passage Five. The number correct on this passage is distributed as follows:

<u>Number of Items Correct</u>	<u>Number of Pupils</u>
4	30
3	22
2	7
1	0
0	0
	<u>59</u>

Fifty-two out of the 59 students, or 88 per cent, who were tested on this passage read Passage Five adequately on the basis of the established criterion. The passage had 88 per cent comprehension at this level of reading ability. No credit was given to those students who scored 2, 1, or 0 items correct. A score of 4 was given no higher weight than a score of 3. These two scores were combined and included to determine the percentage of the total number of pupils who were tested on that passage at that particular reading grade level.

To give further explanation, the distribution of scores for Passage Eight for Grade Equivalent 10 was as follows:

<u>Number of Items Correct</u>	<u>Number of Pupils</u>
4	2
3	9
2	4
1	0
0	1
	<u>16</u>

Eleven students out of 16 students met the criterion of three or four correct items, or 69 per cent of the pupils in reading grade-equivalent 10 were able to understand this passage. This grade equivalent failed to meet the limits of the criterion of 75 per cent comprehension and was considered too difficult for grade-level 10.

Grade-equivalent seven and eight had one and two students, respectively. Because of the small number involved in grade-equivalent seven and eight, neither was used in scaling the difficulty of the paragraphs.

Table VII gives the comprehension percentage for the various paragraphs using the eight different grade equivalents of 9, 10, 11, 12, 13, 14, 15, and 16 and higher.

TABLE VII
COMPREHENSION PERCENTAGES OF READING GRADE LEVELS OF 12
PASSAGES OF GENERAL PSYCHOLOGY TEXTBOOK MATERIALS

Para- graph Numbers	Reading-Grade Equivalents							
	9.0-	10.0-	11.0-	12.0-	13.0-	14.0-	15.0-	16.0+
	9.9	10.9	11.9	12.9	13.9	14.9	15.9	
	n=13	n=16	n=25	n=64	n=96	n=86	n=59	
1	69	81*	84*	81*	90*	92*	93*	94*
2	85**	63	88*	89*	96*	95*	100*	100*
3	92*	94*	92*	97*	98*	98*	98*	97*
4	69	75*	80**	81*	81*	92*	95*	97*
5	77*	63	60	83*	79*	85*	88*	94*
6	77*	69	60	77*	92**	95*	91*	97*
7	62	75**	72	86*	93*	93*	100*	97*
8	62	69	76*	75**	91*	94*	100*	94*
9	31	81*	76*	55	70	79*	71	85**
10	62	81*	76*	89*	98*	98*	100*	100*
11	38	63	68	62	86*	86**	97*	97*
12	46	25	32	42	58	70	78**	85*
* and **	-- Paragraphs within the comprehension of the reading grade levels heading the columns.							
**	-- Paragraphs included in the statistical correlation.							

Statistical Correlation of Obtained Data

The product-moment coefficient of correlation may be thought of essentially as that ratio which expresses the extent to which changes in

in one variable are accompanied by, or are dependent upon, changes in a second variable. The variable of grade-equivalent was correlated with the variable of readability index obtained from passages in the Test of Comprehension. The two sets of variables have been analyzed by means of a product-moment correlation and are illustrated in Table VIII.

TABLE VIII

PRODUCT-MOMENT COEFFICIENT OF CORRELATION OF DATA OBTAINED
FROM CORRELATING GRADE-EQUIVALENT LEVELS OF STUDENTS
WITH READABILITY INDEX OF READING PARAGRAPHS

Average Grade Level X	Readability Index Y	XY	X ²	Y ²
9.5	7.78	73.910	90.24	60.5284
10.6	9.27	98.262	112.36	85.9329
11.4	8.47	96.558	129.96	71.7409
12.5	9.22	115.558	156.25	85.0084
13.4	8.95	119.930	179.56	80.1025
14.4*	9.69	139.536	207.36	93.8961
15.4*	9.89	152.306	237.16	97.8121
16.6*	9.37	155.542	275.56	87.7969
$\Sigma X=103.8$ N = 8	$\Sigma Y=72.64$	$\Sigma XY=951.294$	$\Sigma X^2=1388.46$	$\Sigma Y^2= 662.8182$
$r_{xy} = .7558$				
* Extrapolated				

This study was designed to test the experimental hypothesis that no significant difference exists between the coefficient of correlation obtained in the original cross-validation of the Dale-Chall Readability

Formula and the coefficient of correlation obtained by correlating the readability index of a comprehension test with the reading-grade level of certain students enrolled in general psychology at Oklahoma State University who are able to answer three questions out of four on the test of comprehension. The coefficient of correlation and the mathematical derivation of this product-moment correlation have been presented in Table VIII. A coefficient of .7558 was obtained. In the study in which the Dale-Chall Formula was reported (1948), these authors reported a coefficient of .90 in arriving at a cross-validation index.

To test significance of the difference between two \underline{r} 's requires that the \underline{r} 's be converted into Fisher's \underline{z} function (Garrett, 1958). The formula for the Standard Error of the difference between two \underline{z} 's is

$$S. E. D_{\underline{z}} = S. E. \underline{z}_1 - \underline{z}_2 = \sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}$$

where $S. E. \underline{z} = \frac{1}{\sqrt{N-3}}$ and N_1 and N_2 are the sizes of the two samples.

In the study under investigation $N_1 = 19$, $N_2 = 8$, $\underline{z}_1 = 1.47$, and $\underline{z}_2 = 1.00$. From a conversion table it can be found that \underline{r} 's of .90 and .76 correspond to \underline{z} 's of 1.47 and 1.00, respectively. On substituting the data in the formula, it was found that

$$\begin{aligned} S. E. (1.47 - 1.00) &= \sqrt{\frac{1}{19-3} + \frac{1}{8-3}} \\ &= .51 \text{ (to two decimels).} \end{aligned}$$

Dividing .47 (i. e., $1.47 - 1.00$) by .51, a Critical Ratio of .92 is obtained. A score of 2.08 is necessary to reach the .05 level of significance. Since the Critical Ratio score of .92 is well below the

required significance score of 2.08, results from data in this study demonstrate no significant difference. The null hypothesis cannot be rejected. The null hypothesis is accepted. Based on the evidence from this study, the cross-validation coefficient of correlation obtained by Dale and Chall (1948) and the cross-validation coefficient of correlation obtained in this present study do not really differ except by chance.

Summary

Ten sections of General Psychology at Oklahoma State University were selected at random from the 35 sections available during the Fall Semester, 1961, and Spring Semester, 1962. The 396 students in these sections of general psychology were administered The Nelson-Denny Reading Test and a Test of Comprehension constructed from general psychology textbook materials. The population included with the study was students enrolled in General Psychology drawn from six undergraduate colleges and one graduate college. The two tests were administered two weeks apart for nine sections and three weeks apart for one section.

The instrument used to measure reading achievement was The Nelson-Denny Reading Test, 1960 Revision. The 1939 form of this instrument received favorable reviews in Buros' Fifth Mental Measurement Yearbook. The coefficients reported in the manual for the 1960 revision were .830 when correlated with the American Council on Education Test and .730 when correlated with the Co-operative English Test. These coefficients seemed to be adequate for consideration of this instrument as a valid measure of reading-achievement level. The 1939 form had been

used by Chall in preparing her cross-validation study, later reported as one of the criteria in the Dale-Chall Readability Formula.

If the results obtained from this present study were to yield data of value for future research and application in readability, it seemed desirable that the criterion of comprehension be compared with studies reported in literature. Lorge, Flesch, and Chall used a criterion of 75 per cent. Since an effort had been made to design this study similar to Chall's, a criterion of 75 per cent comprehension was also used.

Student reading-achievement levels were correlated with readability predictions of the passages in the Test of Comprehension. Grade-equivalent groups were determined by The Nelson-Denny Reading Test. The inclusion of a grade-equivalent above the criterion of 75 per cent was based on 75 per cent of the students within that grade-equivalent group being able to answer three or four questions correctly on a particular passage. When the readability index of the Comprehension Test was correlated with the reading-achievement level of students participating in this study, a coefficient of correlation of .7558 was obtained. When the coefficient obtained by Dale and Chall in their original cross-validation study was compared for significance to the coefficient obtained in this study, no significant difference existed. Results from data obtained in this present research reveal that when the Dale-Chall Readability Formula is used to evaluate general psychology textbook material used at Oklahoma State University, no significant difference exists between the coefficient obtained when the original cross-validation study was reported and the coefficient obtained for this present study.

CHAPTER V

SUMMARY AND CONCLUSIONS

General Summary of the Investigation

This investigation examined the validity of the Dale-Chall Readability Formula when this formula was used to predict reading-grade level difficulty of general psychology textbook material used at Oklahoma State University. The null hypothesis that no difference existed between the coefficient of correlation obtained in the original cross-validation of the Dale-Chall Readability Formula and the coefficient of correlation obtained from data within this present study was tested. Students enrolled in ten randomly selected sections of General Psychology, five in the Fall Semester, 1961, and five in the Spring Semester, 1962, were included in the sample population. Data from 396 students were included in this present study. A reading test, The Nelson-Denny Reading Test, 1960 Revision, and a Test of Comprehension of general psychology textbook material were administered to each student participating in this study. No less than 14 days nor more than 21 days separated the administration of the reading test and the administration of the Test of Comprehension.

Summary of Test of Comprehension

The Test of Comprehension used to measure comprehension level of each student consisted of 12 passages and 48 questions. Each of the

12 passages was followed by four multiple-choice questions. The instrument used to measure comprehension level was constructed for use in this present study. Item analysis for each separate question used in the Completed Test of Comprehension produced item-validity index scores ranging from .21 to .87, which were above the minimum index criterion for item analysis (Garrett, 1959).

The split-half technique for determining reliability was used. The odd-even split obtained a coefficient of correlation of .7256. Correction of this Pierson r by use of the Spearman-Brown Prophecy Formula revealed a coefficient of .8409. An r of .8409 denotes high to very high relationship (Garrett, 1958). Even though no appropriate criterion was available for correlation with the instrument used in this research, it has been shown statistically that the Test of Comprehension meets requirements of test construction and, to the extent that it measures what it purports to measure (student comprehension of certain passages taken from the textbook used in general psychology at Oklahoma State University), is a valid instrument and supports assumption three made in Chapter I.

Summary of the Reading Test

The reading test, The Nelson-Denny Reading Test, 1960 Revision, is a revised form of the original 1939 form. Reviews by Traxler (1940) and Held (1939) show that the 1939 form had a validity index which compared favorably with other reading tests of that time. The author of the 1960 revision reported validity coefficients of .830 and .730, both high enough to denote high relationship with other tests measuring similar skills. This reading test reported norms identifiable as grade

levels. The data reported here substantiate the use of The Nelson-Denny Reading Test, 1960 Revision, as an adequate criterion to measure reading-achievement level and support assumptions one and two made in Chapter I.

Summary of Results

The reading test separated the 396 students used in this study into 10 grade-equivalent groups, grade 7.0 through grade 16 and higher. Grade-equivalent groups from grade 9.0 through grade 16 and higher were investigated for comprehension level of each passage of the Test of Comprehension. Correlating the readability index of each passage with the reading-grade level of the grade-equivalent group in which 75 per cent of each group were able to answer three or four questions correctly obtained a coefficient of correlation of .7558. A coefficient of .7558 denotes high to very high relationship.

When the null hypothesis was tested, a Critical Ratio score of .92 was obtained. This score was well below the score of 2.08 required for .05 level of significance. The null hypothesis cannot be rejected. No significant difference exists between the two coefficients.

Concluding Statement

The results of this study have been presented as an attempt to aid in the evaluation of readability of general psychology textbook material. No previous studies were found in the literature in which general psychology had been tested. Further investigations that have similar statistical designs as the method used in this investigation and obtaining coefficients that compare favorably with the coefficient

found in this present study will need to be completed before the reader can place too high level of confidence in the use of the Dale-Chall Readability Formula as a grade-level predictor when evaluating general psychology textbook material. Any general conclusions drawn beyond the scope of this present study would be premature and not statistically sound.

SELECTED BIBLIOGRAPHY

- Allen, William. "Readability of Instructional Film Commentary," Journal of Applied Psychology, XXXVI (June, 1952), pp. 164-68.
- Bergman, W. G. "Objective and Subjective Grade Placement of Supplementary Readers," Reconstructing Education Thru Research, Official Report of the American Education Research Association, Washington: American Research Association, 1936, pp. 263-71.
- Betts, Emmett A. "Readability: Its Application in the Elementary Schools," Journal of Educational Research, XLII (February, 1949), pp. 438-59.
- Burk, Cassie. "A Study of the Influence of Some Factors in Style of Composition on the Interest, Comprehension, and Rate of Reading of Fourth Grade Pupils," Journal of Experimental Education, IV (June, 1936), pp. 303-52.
- Bushwell, G. T. How Adults Read. Chicago: University of Chicago Press, 1937.
- Chall, Jeanne S. "This Business of Readability," Educational Research Bulletin, XXVI (January 15, 1947), pp. 1-13.
- _____. "Graded Reading Paragraphs in Health Education." (unpub. Master's thesis, Ohio State University, 1947).
- _____, and Dial, Harold E. "Predicting Listener Understanding and Interest in Newscasts," Educational Research Bulletin, XXVII (September 15, 1948), pp. 141-53, 168.
- _____. Readability: An Appraisal of Research and Application. Columbus: Ohio State University Press, 1958.
- Cronbach, Lee J. Essentials of Psychological Testing. New York: Harper and Brothers, 1949.
- Dale, Edgar, and Tyler, Ralph W. "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability," Library Quarterly, IV (July, 1934), pp. 384-412.
- _____, and Chall, Jeanne S. "A Formula for Predicting Readability," Educational Research Bulletin, XXVII (January 21, 1948).

- _____, and Chall, Jeanne S. "The Concept of Readability," Elementary English, XXVI (January, 1949), pp. 19-26.
- _____, and Chall, Jeanne S. "Techniques for Selecting and Writing Readable Materials," Elementary English, XXV (May, 1949), pp. 250-58.
- DeLong, Vaughn R. "Primary Promotion By Reading Levels," Elementary School Journal, XXXVIII (May, 1938), pp. 663-71.
- Dolch, Edward W. "Vocabulary Burden," Journal of Educational Research, XVII (March, 1928), pp. 170-83.
- _____. Problems in Reading. Champaign: The Garrard Press, 1948, pp. 229-55.
- _____. "The Use of Vocabulary Lists in Predicting Readability and in Developing Reading Materials," Elementary English, XXVI (March, 1949), pp. 142-49.
- Dunlap, Carolyn C. "Readability of Newspaper Items and of Basic Reading Material," Elementary School Journal, LI (May, 1951), pp. 499-501.
- _____. "Readability Measurement: A Review and Comparison." (unpub. Ph. D. dissertation, University of Maryland, 1954).
- Elliott, Catherine J. "A Critical Analysis of the Objective Method of Measuring Reading Difficulty," Pittsburg Schools, XV (May-June, 1941), pp. 201-09.
- Engleman, Finis E. "The Relative Merits of Two Forms of Discourse When Applied to Children's Factual Content Reading Material," Journal of Educational Research, XXIX (March, 1936), pp. 524-31.
- Faison, Edmund W. J. "Readability of Children's Textbooks," Journal of Educational Psychology, XLII (January, 1951), pp. 43-51.
- Farr, James N., and Jenkins, James J. "Tables For Use With the Flesch Readability Formulas," Journal of Applied Psychology, XXXIII (June, 1949), pp. 275-78.
- _____; Jenkins, James J.; and Paterson, Donald G. "Simplification of Flesch Reading Ease Formula," Journal of Applied Psychology, XXXV (October, 1951), pp. 333-37.
- _____, et. al. "Reply to Klare and Flesch re 'Simplification of Flesch Reading Ease Formula'," Journal of Applied Psychology, XXXVI (February, 1952), pp. 55-57.

Figurel, J. Allen. "Relative Difficulty of Reading Material for Ninth Grade Literature," Pittsburg Schools, XVI (January-February, 1942), pp. 125-38.

Flanagan, J. C. "General Consideration in the Selections of Test Items," Journal of Educational Psychology, XXX (December, 1939), pp. 674-80.

Flesch, Rudolph. "Estimating the Comprehension Difficulty of Magazine Articles," Journal of General Psychology, XXVIII (January, 1943), pp. 63-80.

_____. Marks of Readable Style. Teacher College Contributions to Education No. 897, New York: Bureau of Publications, Teachers College, Columbia University, 1943.

_____. "A New Readability Yardstick," Journal of Applied Psychology, XXXII (June, 1948), pp. 221-33.

_____. "New Facts About Readability," College English, X (January, 1949), pp. 225-26.

_____. "How to Test Readability: Measuring the Level of Abstraction," Journal of Applied Psychology, XXXIV (December, 1950), pp. 384-90.

_____. "Reply to 'Simplification of Flesch Reading Ease Formula'," Journal of Applied Psychology, XXXVI (February, 1952), pp. 54-55.

Forbes, Fritz W., and Cottle, William C. "A Method For Determining Readability of Standardized Tests," Journal of Applied Psychology, XXXVII (June, 1953), pp. 185-90.

Garrett, Henry E. Statistics in Psychology and Education. New York: Longmans, Green and Company, 1947.

_____. Statistics in Psychology and Education. New York: Longmans, Green and Company, 1958.

_____. Testing for Teachers. New York: American Book Company, 1959.

Gates, Arthur I. Interest and Ability in Reading. New York: The Macmillan Company, 1931.

Gilinsky, Alberta S. "How Valid is the Flesch Readability Formula?" American Psychologist, III (July, 1948), p. 261.

Gray, William S., and Leary, Bernice W. What Makes a Book Readable. Chicago: University of Chicago Press, 1935.

- _____. "Progress in the Study of Readability." Elementary School Journal, XLVII (May, 1947), pp. 490-99.
- Guilford, J. P. "The Phi-coefficient and Chi-square as Indices of Item Validity." Psychometrika, VI (February, 1941), pp. 11-19.
- Gunning, Robert. The Technique of Clear Writing. New York: McGraw-Hill Company, 1952.
- Haskes, H. E.; Lindquist, E. R.; and Mann, C. R. Achievement Examinations. Boston: Houghton Mifflin Company, 1936, pp. 17-160.
- Hayes, Patricia M.; Jenkins, James J.; and Walker, Bradley J. "Reliability of The Flesch Readability Formula." Journal of Applied Psychology, XXXIV (February, 1950), pp. 22-26.
- Held, Omar C. "Nelson-Denny Reading Tests as an English Placement Test." School and Society, IL (January, 1939), pp. 318-21.
- Holzinger, Karl J. Statistical Methods for Students in Education. Boston: Ginn and Company, 1928, p. 278.
- Hotchkiss, Sanford N., and Paterson, Donald G. "Flesch Readability Reading List." Personnel Psychology, III (Autumn, 1950), pp. 327-44.
- Jackman, Mabel E. "The Relationship Between Maturity of Content and Simplicity of Style in Selected Books of Fiction." Library Quarterly, XI (July, 1941), pp. 302-27.
- Johnson, George R. "An Objective Method of Determining Reading Difficulty." Journal of Educational Research, XXI (April, 1930), pp. 283-87.
- Johnson, Mary E. "The Vocabulary Difficulty of Control Subjects in Grade Five." Elementary English, XXIX (May, 1952), pp. 277-80.
- Kearl, Bryant. "A Closer Look at Readability Formulas." Journalism Quarterly, XXV (December, 1948), pp. 344-48.
- Keboch, F. D. "Variability of Word Difficulty in Five American History Textbooks." Journal of Educational Research, XV (January, 1927), pp. 22-26.
- Klare, George R. "Measures of The Readability of Written Communication: An Evaluation." Journal of Educational Psychology, XXXV (June, 1951), pp. 151-56.
- _____. "A Table For Rapid Determination of Dale-Chall Readability Scores." Educational Research Bulletin, XXXI (February 13, 1952), pp. 43-47.

- _____. "A Note on 'Simplification of Flesch Reading Ease Formula'." Journal of Applied Psychology, XXXVI (February, 1952), p. 53.
- Kuder, G. F., and Richardson, M. W. "The Theory of The Estimation of Test Reliability." Psychometrika, II (September, 1937), pp. 151-60.
- Latimer, Edward H. "A Comparative Study of Recent Techniques for Studying Readability." University of Pittsburg Bulletin, 1948, pp. 246-56.
- Leary, Bernice E. "Difficulties in Reading Material." Reading in General Education, ed. William S. Gray (American Council on Education, 1940), pp. 272-306.
- Levy, Jerome. "Readability Level and Differential Test Performance: A Language Revision of the Study of Values." Journal of Educational Psychology, XLIX (February, 1958), pp. 6-12.
- Lewerenz, Alfred S. "Measurement of the Difficulty of Reading Material." Educational Research Bulletin, Los Angeles City Schools, VIII (March, 1929), pp. 11-16.
- Lindquist, E. F., and Cook, Walter W. "Experimental Procedures in Test Evaluation." Journal of Experimental Education, I (March, 1933), pp. 163-85.
- Lively, Bertha A., and Pressey, S. L. "A Method For Measuring the 'Vocabulary Burden' of Textbooks." Educational Administration and Supervision, IX (October, 1923), pp. 389-98.
- Lorge, Irving. "Predicting Reading Difficulty of Selections For Children." Elementary English Review, XVI (April, 1939), pp. 229-33.
- _____. "Predicting Readability." Teachers College Record, XLV (March, 1944), pp. 404-19.
- _____. "The Lorge and Flesch Readability Formulae: A Correction." School and Society, LXVII (February 21, 1948), pp. 141-42.
- _____. "Readability Formulae: An Evaluation." Elementary English, XXVI (February, 1949), pp. 86-95.
- _____, and Kruglov, Lorraine. "The Relationship Between the Readability of Pupils' Compositions and Their Measured Intelligence." Journal of Educational Research, XLIII (February, 1950), pp. 467-74.
- _____. "The Relationship Between Merit of Written Expression and Intelligence." Journal of Educational Research, XLIV (March, 1951), pp. 507-19.

- Ludwig, Merritt C. "Hard Words and Human Interest: Their Effects on Readership." Journalism Quarterly, XXVI (June, 1949), pp. 167-71.
- Lyman, H. B. "Flesch Count and Readership of Articles in a Midwestern Farm Paper." Journal of Applied Psychology, XXXIII (February, 1949), pp. 78-80.
- Michaelis, John U., and Tyler, Fred T. "A Comparison of Reading Ability and Readability." Journal of Educational Psychology, XLII (December, 1951), pp. 49-98.
- Miller, L. R. "Reading-Grade Placement of the First Twenty-Three Books Awarded the John Newbery Prizes." Elementary School Journal, XLVI (March, 1946), pp. 394-400.
- Miller, R. A. "Relation of Reading Characteristics to Social Indexes." American Journal of Sociology, XI (1936), pp. 738-51.
- Morgan, Clifford T. Introduction to Psychology. New York: McGraw-Hill, 1961.
- Morris, Elizabeth C., and Holverson, Dorothy. "Idea Analysis Technique." (unpub. manuscript, Teachers College, Columbia University, 1938).
- Murphy, Donald R. "How Plain Talk Increases Readership 45% to 66%." Printers Ink, CCXX (September 19, 1947), pp. 35-37.
- McClusky, Howard Y. "A Quantitative Analysis of the Difficulty of Reading Materials." Journal of Educational Research, XXVIII (December, 1934), pp. 276-82.
- McCracken, Robert A. "An Experiment With Contrived Readability in Fifth and Sixth Grades." Journal of Educational Research, LII (March, 1959), pp. 277-78.
- Nelson, M. J., and Denny, E. C. Examiners Manual: The Nelson-Denny Reading Test. (Revised by James I. Brown), Boston: Houghton Mifflin Company, 1960.
- Odgen, C. K., and Richards, I. A. The Meaning of Meaning. New York: Harcourt, Brace and Company, 1927, pp. 185-208.
- Ojeman, Ralph. "The Reading Ability of Parents and Factors Associated With Reading Difficulty of Parent Education Materials." Research in Parent Education, II, ed. George D. Stoddard, Iowa City: University of Iowa, 1934.

- Painter, Helen Welch, and Franzen, Carl G. F. "A Synthesis of Research on the Placement of Reading Material in Secondary School Literature." Journal of Educational Research, XXXIX (December, 1945), pp. 304-06.
- Patty, W. W., and Painter, W. I. "Improving Our Method of Selecting High-School Textbooks." Journal of Educational Research, XXIV (June, 1931), pp. 23-32.
- Richardson, M. W., and Adkins, D. C. "A Rapid Method of Selecting Test Items." Journal of Educational Psychology, XXX (December, 1939), pp. 681-87.
- _____, and Kuder, G. F. "The Calculation of Test Reliability Coefficients Based on the Method of Rational Equivalence." Journal of Educational Psychology, XXX (December, 1939), pp. 681-87.
- Robinson, Francis P. "Comprehension Difficulty and Inspirational Value." Pedagogical Seminary and Journal of Genetic Psychology, LVI (March, 1940), pp. 53-65.
- _____. "The Effect of Language Style on Reading Performance." Journal of Educational Psychology, XXXVIII (March, 1947), pp. 149-56.
- Russell, David H., and Merrill, Anne F. "Children's Librarians Rate the Difficulty of Well-Known Juvenile Books." Elementary English, XXVIII (May, 1951), pp. 263-68.
- _____, and Fea, Henry R. "Validity of Six Readability Formulas as Measures of Juvenile Fiction." Elementary School Journal, LII (November, 1961), pp. 136-44.
- Smith, Edgar A. "Devereux Readability Index." Journal of Educational Research, LIV (April, 1961), pp. 298-303.
- Smith, Ruth I. "An Investigation of the Readability of Recently Published History and Geography Textbooks and Related Materials for the Fourth Grade." University of Pittsburg Bulletin, XLVIII (June 5, 1952), pp. 521-27.
- Spache, George. "A New Readability Formula for Primary-Grade Reading Materials." Elementary School Journal, LIII (March, 1953), pp. 410-13.
- Spaulding, Seth. "Two Formulas for Estimating the Reading Difficulty of Spanish." Educational Research Bulletin, XXX (May, 1951), pp. 117-24.

Stadtlander, Elizabeth. "A Scale for Determining and Evaluating Reading Materials for the Middle Grades." University of Pittsburg Bulletin, XV (1939), pp. 347-54.

Stolurow, L. M., and Newman, J. R. "A Fractional Analysis of Objective Features of Printed Language Presumably Related to Reading Difficulty." Journal of Educational Research, LII (March, 1959), pp. 243-51.

Stone, Clarence R. Better Primary Reading. St. Louis: Webster Publishing Company, 1936.

_____. "Measures of Simplicity and Beginning Texts in Reading." Journal of Educational Research, XXXI (February, 1938), pp. 447-50.

Studies in Item Analysis and Prediction. ed. Herbert Solomon, Stanford: Stanford University Press, 1960.

Swanson, Charles E. "Readability and Readership: Controlled Experiment." Journalism Quarterly, XXV (December, 1948), pp. 339-41.

_____, and Fox, Harland G. "Validity of Readability Formulas." Journal of Applied Psychology, XXXVII (April, 1953), pp. 114-18.

Taylor, H. K. "What Magazines Do Book Readers Read?" Publishers Weekly, GXXXIII (1938), pp. 1922-26.

Thorndike, Edward L. "Improving the Ability to Read." Teachers College Record, XXXVI (October, November, December, 1934), pp. 1-19.

Traxler, Arthur E. "The Measurement and Improvement of Silent Reading at the Junior-High-School Level." (unpub. Ph.D. dissertation, University of Chicago, 1932).

_____. "One Reading Test Serves the Purpose." Clearing House, XIV (March, 1940), pp. 419-21.

Twedt, D. W. "A Table for Use With Flesch's Level of Abstraction Readability Formula." Journal of Applied Psychology, XXXV (June, 1951), pp. 157-59.

Tyler, Ralph W. "Ability to Use Scientific Method." Educational Research Bulletin, XI (January 6, 1932), pp. 1-9.

Vogel, Mabel, and Washburne, Carleton. "An Objective Method of Determining Grade Placement of Children's Reading Material." Elementary School Journal, XXVIII (January, 1928), pp. 373-81.

- Walther, Cyrilla. "The Reading Difficulty of Magazines," School Review, LI (February, 1943), pp. 100-05.
- Washburne, Carleton, and Vogel, Mabel. Winnetka Grade Book List. Chicago: American Library Association, 1926.
- _____, and Morphett, Mabel V. "Grade Placement of Children's Books," Elementary School Journal, XXXVIII (January, 1938), pp. 355-64.
- Wert, James E. "A Technique for Determining Levels of Group Reading." Education Research Bulletin, XVI (May 19, 1937), pp. 113-21.
- _____, Neidt, Charles O.; and Ahmann, J. Stanley. Statistical Methods in Educational and Psychological Research. New York: Appleton-Century-Crofts, Inc., 1954, p. 334.
- Wheeler, L. R., and Wheeler, V. D. "Selecting Appropriate Reading Materials." Elementary English, XXV (December, 1948), pp. 478-89.
- Wilson, Mary C. "The Effect of Amplifying Material upon Comprehension." Journal of Experimental Education, XIII (September, 1948), pp. 5-8.
- Witherington, H. Carl. "Readability of Textbooks in Educational Psychology." Journal of Educational Research, XLVI (November, 1952), pp. 227-30.
- Yoakam, Gerald A. "A Technique for Determining the Difficulty of Reading Materials." (unpub. Ph.D. dissertation, University of Pittsburg, 1939).
- _____. "The Reading Difficulty of School Textbooks," Elementary English Review, XXII (December, 1945), pp. 304-09.
- _____. Basal Reading Instruction. New York: McGraw-Hill Book Company, 1955, pp. 329-40.
- Young, James R. "Understanding Radio News: The Effect of Style." Journalism Quarterly, XXVII (March, 1950), pp. 19-23.

APPENDIX A

PRELIMINARY TEST OF COMPREHENSION ADMINISTERED
TO DETERMINE ITEM ANALYSIS

TEST OF COMPREHENSION

Directions to students: This is a test to determine how well you are able to read selected passages and answer questions that are asked about the passages.

There are twelve reading passages in this test. Read a selection through completely, then answer the questions following it. When you have completed one passage, go to the next. Keep working until you have completed all twelve passages or until you are told to stop. You may look back at the material you have read, if you wish, in order to answer the questions correctly, but do not puzzle too long over any one question. Pass on to the next after a reasonable effort.

Mark all answers on the answer sheet.

PASSAGE ONE

It is often said that man has five senses; sight, hearing, smell, taste, and touch. These five, however, are only the obvious ones we know we have from everyday experience. In addition, there are several more that have been discovered by examining the organs of the body and by experimenting with their function. The fifth of the common five is a skin touch. There are also sense organs in the muscles, tendons, and joints that afford a feeling of pressure within the body. These organs are called kinesthetic receptors, and the sense they serve is kinesthesia. There are also, finally, other sense organs in the head associated with the hearing apparatus. These are called vestibular receptors; they respond to the force of gravity and to rotation of the head. They are the key sense organs in our sense of balance. The real count then is something like ten senses, rather than five. Some of these can be further subdivided into other senses.

The sensitive elements of each sense organ are called receptors. A receptor is a cell that is specialized to respond to relatively small changes of a particular kind of energy. Some receptors, those for sight and smell, for example, are really nerve cells that migrated out from the brain and became specialized for their particular function.

Each of these responds primarily to a particular kind of physical energy. Smell and taste are chemical senses, for they respond to chemical energy.

1. This passage is concerned with explaining the:
 - a. common five senses.
 - b. ten senses.
 - c. sense organs.
 - d. kinesthetic receptors.
2. Vestibular receptors respond to:
 - a. electromagnetic energy.
 - b. chemical energy.
 - c. force of gravity.
 - d. thermal energy.
3. The sense organs in the muscles, tendons, and joints are known as:
 - a. vestibular receptors.
 - b. kinesthetic receptors.
 - c. unspecialized receptors.
 - d. none of these.
4. The number of senses whose receptors are associated with skin cells are:
 - a. one.
 - b. two.
 - c. three.
 - d. four.

(Go on to next page)

5. Balance is maintained through the:
 - a. vestibular receptors.
 - b. kinesthetic receptors.
 - c. ten senses.
 - d. skin cells.
6. It is now believed that the number of senses in the human body is:
 - a. five.
 - b. ten.
 - c. something like ten.
 - d. more than ten.
7. The nerve cells that migrated from the brain and become specialized are:
 - a. sight and smell.
 - b. pain.
 - c. taste.
 - d. hearing.
8. There are _____ mechanical senses.
 - a. ten.
 - b. four.
 - c. two.
 - d. five.

PASSAGE TWO

The inner ear is by far the most complicated of the three major parts of the ear. It consists of two kinds of sense organs, one concerned in the sense of balance, and the other in hearing. The organs for balance are called the vestibular sense organs and will be discussed later in this chapter. Because the structure for hearing is spiraled like a snail, it is called the cochlea, meaning snail shell. The cochlea has three different ducts or canals spiraling around together, separated from each other by membranes. Each duct is filled with fluid. Sound vibrations enter them from the ossicles of the middle ear through a membrane known as the oval window, located at the end of the vestibular canal, but too small to show in a drawing. Thus when a sound wave moves the ossicles back and forth, this movement is transmitted to the fluid of the cochlea.

The important event that takes place in the cochlea is the stimulation of nerve cells -- called hair cells because they have hairs on their ends -- located in the organ of Corti on the basilar membrane separating two of the canals. Pressure changes of the fluid in the canals stimulate these hair cells, causing nerve impulses to be generated and to be conducted through the auditory nerve to the brain.

The basilar membrane beneath the hair cells varies in width from one end of the cochlea to the other, and its width correlates with the frequencies of sounds to which the hair cells are most sensitive. It is narrowest at the end toward the oval window and becomes progressively wider toward the other end.

9. This passage is most concerned with explaining:
 - a. sound vibrations.
 - b. stimulation of nerve cells.
 - c. the structure of the cochlea.
 - d. the inner ear.

(Go on to next page)

10. The most complicated part of the ear is the:
 - a. outer ear.
 - b. inner ear.
 - c. middle ear.
 - d. area between the middle ear and the inner ear.
11. The vestibular sense organs are:
 - a. the organs used for hearing.
 - b. contained in a bony structure.
 - c. contained in a spiral structure.
 - d. the organs of balance.
12. The cochlea has three different ducts or canals spiraling around together, separated by:
 - a. bones.
 - b. fluid.
 - c. membranes.
 - d. nerve endings.
13. When a sound wave moves the ossicles back and forth:
 - a. this movement is transmitted to the fluid of the cochlea.
 - b. this movement is transmitted to wax in the middle ear.
 - c. the cochlea stimulates the fluid within it.
 - d. the cochlea is sensitized.
14. Nerve impulses are generated when:
 - a. sound strikes the outer ear.
 - b. a change of pressure in the fluid of the canal occurs.
 - c. both a and b occur.
 - d. neither a nor b occur.
15. The most important event that takes place in the cochlea is:
 - a. the stimulation of nerve cells.
 - b. a change of fluid pressure.
 - c. generation of auditory sounds.
 - d. the brain receives a nerve impulse.
16. The membrane beneath the hair cells that vary in width from one end of the cochlea to the other is known as:
 - a. the vestibular sense organ.
 - b. the basilar membrane.
 - c. the ossicles membrane.
 - d. the oval window.

PASSAGE THREE

Women have a reputation for their ability to talk, and they deserve it. Girls, on the average, learn to talk at an earlier age than boys do. Later, when they can take intelligence tests, they do better on all tests involving the use of language. Thus, they are superior to men in verbal abilities. Closely related to this is the fact that they generally do better on items involving social relations.

Girls are also better on two other related kinds of abilities: perceiving details quickly and accurately and making quick, accurate manual movements. This may account, in part, for the fact that most women are employed in business and industrial work in clerical and secretarial jobs and also for the good record they made during the Second World War in many kinds of industrial jobs. Women also seem to have better immediate memories than men. You will recall that the Stanford-Binet has a series of numbers or words. On the average, women are better in these items than are men.

(Go on to next page)

On the Stanford-Binet test, boys surpass girls on the items that involve spatial, numerical, and mechanical tasks. Correlated with this is the fact that they do better on tasks involving the perception of spatial relationships, the comprehension of mechanical tools and machines, and mathematical ability. Boys are better in working with numbers and in numerical-reasoning tasks.

We see, then, that men and women differ significantly in certain specific abilities. These conclusions, reached from careful analysis of tests of ability, fit in pretty well with popular conceptions of the difference in ability between men and women.

17. This passage is concerned primarily with:
 - a. women's ability to talk.
 - b. the fact that women score higher on some items of the Stanford-Binet test than do men.
 - c. pointing out that men and women differ in certain abilities.
 - d. the fact that girls perceive details quickly and accurately.
18. On the Stanford-Binet test, boys surpass girls in:
 - a. items that involve spatial tasks.
 - b. items that involve numerical tasks.
 - c. items that involve mechanical tasks.
 - d. all the above items.
19. Boys do better than girls on tasks that:
 - a. involve spatial relationships.
 - b. involve accurate manual movements.
 - c. involve verbal abilities.
 - d. involve social relationships.
20. On tasks that require immediate memory:
 - a. boys do better than girls.
 - b. girls do better than boys.
 - c. there is no difference between boys and girls.
 - d. it has not been determined whether boys or girls do better.
21. When girls take intelligence tests, they do better than boys on tests involving:
 - a. spatial tasks.
 - b. mathematical ability.
 - c. numerical reasoning.
 - d. the use of language.
22. Girls surpass boys in:
 - a. comprehension of mechanical tools and machines.
 - b. mathematical ability.
 - c. perceiving details quickly and accurately.
 - d. mechanical tasks.
23. Girls, being better than boys in perceiving details quickly and accurately and making quick accurate movements, are able to fill successfully such jobs as:
 - a. clerical tasks.
 - b. mechanical tasks.
 - c. mathematical tasks.
 - d. spatial tasks.
24. The passage leaves little doubt that:
 - a. boys are better than girls in all tasks.
 - b. boys are better than girls in some tasks.
 - c. boys are inferior to girls in all tasks.
 - d. there is no measurable difference between boys and girls in any task.

(Go on to next page)

PASSAGE FOUR

The maze has long been a favorite instrument of psychologists for studying the learning process. A maze is any apparatus consisting of a series of paths and choice points between a starting position and the goal box or outlet. Mazes for animals usually consist of alleys arranged in a pattern of Ys or Ts through which the animal swims, walks, or runs.

As an instrument for studying animal learning, the maze has one unique feature. It provides the opportunity to study serial learning; that is, learning a whole series of responses, not just one, leading up to a goal. This raises some interesting questions and throws some light on the processes involved in problem solving. Otherwise, the maze requires instrumental learning just as a Skinner box does, but it also entails perceptual learning. Thus the maze combines serial learning with two varieties of learning we have already studied. By analyzing the processes concerned in maze learning, we gain some understanding of the mechanisms of complex trial-and-error learning such as are involved in human thinking and problem solving.

The simplest possible way in which the learning of a maze might take place is that each response at a choice point (a left turn or a right turn) might become associated with the response just preceding it.

25. This passage is concerned with explaining:
 - a. animal ways.
 - b. learning theory.
 - c. instrumental learning.
 - d. mazes.
26. The unique feature provided by the maze is:
 - a. the opportunity to study animal learning.
 - b. the opportunity to watch animals run, swim, or walk.
 - c. the opportunity to study instrumental learning.
 - d. the opportunity to study serial learning.
27. Mazes usually consist of alleys arranged in a pattern of:
 - a. Ts
 - b. Ys
 - c. Vs
 - d. a or b.
28. Serial learning is:
 - a. learning a response.
 - b. learning a response after the initial response.
 - c. learning three responses.
 - d. learning a series of responses.
29. The maze combines _____ types of learning.
 - a. two
 - b. three
 - c. four
 - d. five
30. Any apparatus consisting of a series of paths and choice points between a starting position and a goal or outlet is:
 - a. a maze
 - b. a Skinner box.
 - c. a device to measure learning.
 - d. none of these.

(Go on to next page)

31. Maze learning involves a response being associated with:
- the response just preceding it.
 - the two responses preceding it.
 - the three responses preceding it.
 - none of these.
32. Psychologists consider the maze a favorite _____ for studying the learning process.
- device
 - mechanism
 - instrument
 - box

PASSAGE FIVE

Groups of which we are members influence our behavior in various ways, but probably the strongest and most pervasive of the ways are group norms. A norm, as the term implies, is a standard of behavior, but it is more than that. To understand exactly what it is we must refer again to the concepts of role and status.

A role is the behavior expected of us in a particular status. The accent should now be placed on the word "expected." A group can expect certain behavior from us because it confers its disapproval on us if we do not do what is expected. Since most of us acquire a need for social approval, and hence do not wish to incur disapproval, we do what is expected of us. In a word, we conform to our group's expectations.

Group norms seem to emerge, like statuses and social structure, wherever a group is formed. A group exists whenever there is any interaction between individuals. By interaction we mean any conversation, exchange of goods and services, or any joint efforts which could tend to cast group members into any kind of status. The longer people interact, and the more they interact, the more they tend to adopt common ways of interpreting the world and common standards for the behavior of each group member. In other words, group norms to which individuals feel a pressure to conform develop whenever there is any kind of continued interaction among people.

33. A standard of behavior is:
- a mean.
 - a norm.
 - an average.
 - a median.
34. Before a group can exist there first must be:
- interest of the individuals.
 - conformity by some of the individuals.
 - interaction between individuals.
 - common behavior.
35. The expected behavior of a person in a particular status is known as:
- role.
 - conformity to group standards.
 - emerging leadership.
 - seeking status.

(Go on to next page)

36. Group norms develop only when there is:
- conversation.
 - interaction between individuals.
 - common interpretation in communication.
 - continued interaction among people.
37. A group can expect certain behavior from its members because:
- it shows disapproval when members disobey rules.
 - it sets up norms for the group.
 - it rewards the group.
 - none of these.
38. A need for social approval for most of us is:
- acquired.
 - inherited.
 - determined by parents.
 - determined by teachers.
39. Group norms:
- are determined by society.
 - are determined by group leaders.
 - emerge.
 - none of these.
40. Because we wish to be liked by our group:
- we conform to our group's expectations.
 - we interact between individuals.
 - we exchange goods or services.
 - we propose new types of behavior.

PASSAGE SIX

From psychological studies of the nervous system and of bodily changes in emotion, we know that the changes that occur are initiated by a part of the nervous system called the autonomic system. The changes are therefore called autonomic changes.

The autonomic system consists of many nerves leading from the brain and spinal column out to the various organs of the body, including particularly the blood vessels serving both the interior and exterior muscles. The autonomic system has two parts which usually work in opposition to each other. One part, the sympathetic system, increases the heart rate and blood pressure and distributes blood to the exterior muscles. It is this part that comes into play when we become emotional--or at least when we become fearful or angry. The other part of the system is called the parasympathetic system. It tends to be active when we are calm and relaxed. It does many things that, taken together, build up and conserve the body's stores of energy. For example, it decreases the heart rate, reduces blood pressure, and diverts blood to the digestive tract.

When the sympathetic part of the autonomic system steps up its discharges, as it does in emotion, it produces several symptoms that are worth noting. One set of symptoms concerns the circulation of blood. The blood vessels serving the stomach, intestines, and interior of the body tend to contract in emotion, while those serving the exterior muscles of the trunk and limbs tend to become larger. In this way, blood is diverted from digestive functions to muscular

(Go on to next page)

functions, and the body is thus prepared for action that may involve great muscular activity. At the same time, nervous impulses to the heart make it beat harder and faster, which means that the blood pressure goes up and the pulse rate is quickened. Thus more blood is pumped through the circulatory system to the muscles.

Besides producing changes in circulation, the sympathetic system produces several other bodily changes in emotion. Perhaps, when you have been afraid, you have felt some of them yourself. One is a change of breathing.

41. This passage is involved with explaining the function of the:
 - a. central nervous system.
 - b. autonomic nervous system.
 - c. sympathetic nervous system.
 - d. parasympathetic nervous system.
42. Bodily changes occurring during emotional reactions are initiated by the:
 - a. central nervous system.
 - b. autonomic nervous system.
 - c. endocrine glands.
 - d. spinal column.
43. The nervous system(s) responsible for increased heart rate is:
 - a. the sympathetic system.
 - b. parasympathetic system.
 - c. the central nervous system.
 - d. both a and b.
44. The nervous system(s) responsible for build up and conservation of the body's stores of energy is the:
 - a. parasympathetic system.
 - b. sympathetic system.
 - c. central nervous system.
 - d. both b and c.
45. The nervous system(s) in charge of body functions when the body is at rest is the:
 - a. sympathetic system.
 - b. parasympathetic system.
 - c. central nervous system.
 - d. none of these.
46. A symptom of the sympathetic system is:
 - a. reduced blood pressure.
 - b. decreased heart rate.
 - c. decrease in size of exterior muscles.
 - d. contraction of blood vessels in the stomach.
47. When the sympathetic part of the autonomic nervous system steps up its discharges, it:
 - a. causes an increase in blood pressure.
 - b. causes more blood to flow to the stomach.
 - c. causes less blood to be pumped to the muscles.
 - d. all of above.

(Go on to next page)

48. The nervous system consisting of two opposing parts is:
- a. the autonomic nervous system.
 - b. the central nervous system.
 - c. the sympathetic nervous system.
 - d. the parasympathetic nervous system.

PASSAGE SEVEN

In order to understand physiological drives we should know the meaning of a concept called homeostasis. This is the tendency of the body to maintain a balance among internal physiological conditions. Such a balance is essential for the individual's survival. Body temperature must not get too high or too low. Blood pressure must not rise or fall beyond certain limits. The blood must not get too acidic or alkaline; it must not contain too much carbon dioxide; it must not become too concentrated; it must have a certain amount of sugar in it. If these limits are exceeded, the individual becomes sick and he may die. Physiologists have discovered that many homeostatic mechanisms are involved in keeping conditions within normal limits. Consider, for example, the control of body temperature. Normal body temperature usually stays near that point because the body can cool or heat itself. If a person's body temperature tends to get too high, he perspires and the resultant evaporation of liquid cools the body. If his temperature threatens to fall, he shivers and steps up his metabolism. Shivering burns the body's fuel faster and thus generates extra heat. In addition, many animals can insulate themselves against heat loss by fluffing their fur and creating a dead-air space around their skin. All that is left of this mechanism in human beings, however, is the goose pimples one gets when he is too cold.

Physiological mechanisms take care of many of the problems of maintaining a homeostatic balance, but the body also makes use of regulatory behavior--behavior that has the effect of regulating internal physiological conditions--to maintain or restore the balance. Such regulatory behavior is behavior that is instrumental in satisfying physiological needs. When the body becomes depleted of water or food, for example, it cannot maintain a balance by calling on its physiological mechanisms. Rather it must obtain more water and food from the outside. It does this through motivated behavior that normally succeeds in procuring more water and food; after that, the homeostatic balance is restored. The important point, then, is that physiological mechanism for maintaining homeostatic balance within the body.

49. This passage is concerned with the explanation of:
- a. physiological needs.
 - b. internal physiological conditions.
 - c. balance among internal physical conditions.
 - d. metabolism.

(Go on to next page)

50. When the limits of blood sugar are exceeded, the person:
- becomes sick.
 - has excess energy.
 - stores unused energy.
 - is short of energy.
51. Homeostasis may best be explained as:
- control of body temperature.
 - the tendency of the body to maintain a balance among internal physiological conditions.
 - physiological mechanisms to take care of the problems of depletion of food and water.
 - motivated behavior.
52. Physiological mechanisms take care of the problems of maintaining:
- motivated behavior.
 - homeostatic balance.
 - individual survival.
 - both b and c.
53. Regulatory behavior is:
- not motivated by a desire to be comfortable.
 - instrumental in satisfying physiological needs.
 - instrumental in satisfying psychological needs.
 - instrumental in satisfying emotional needs.
54. Cooling takes place when:
- a body perspires.
 - a body shivers.
 - perspiration evaporates.
 - body fuels burn faster.
55. Insulation against heat loss is accomplished by:
- burrowing in the ground by animals.
 - fluffing hair.
 - insulating nests.
 - all of these.
56. Body fuels are burned faster when a person:
- shivers.
 - rests.
 - perspires.
 - none of these.

PASSAGE EIGHT

For a long time it was thought that neurosis and psychosis were two completely different kinds of disorders. "Once a neurotic, never a psychotic" was almost an axiom. By this, it was meant that neurosis and psychosis were such different reactions that they could not develop in the same person. Today we are not so sure about this. Certainly many patients encountered in clinical practice are difficult to classify and appear to be somewhere in between.

In principle, however, we can make a distinction. Whereas the neurotic individual is characterized by anxiety or strenuous defense against anxiety, the psychotic individual is typically one who has lost considerable contact with reality. He may simply withdraw and

(Go on to next page)

fail to respond to things going on around him. Or he may be so excited or depressed that his reactions are quite inappropriate to circumstances. In many instances, his thought processes may be seriously disturbed by delusions and hallucinations. In any case, he tends to live in a world of his own rather than in the real world around him. For this reason, psychosis is more severe than neurosis, and the psychotic person is more likely to require hospitalization and protective care. Most of the inmates of our mental hospitals, particularly those who stay a long time, are psychotic, not neurotic, individuals.

Psychotic reactions are classified into two categories: functional and organic. The functional psychoses are generally considered psychological in origin, though some of the data on the inheritance of functional psychoses raise doubts about this. The organic psychoses are those known to be caused by organic agents, such as the syphilis germ, degenerative changes in old age, alcoholism, and the like. As psychologists, we have a greater interest in the functional psychoses. These may be described under three main headings: affective reaction, paranoid reactions, and schizophrenic reactions.

57. Functional psychoses may be described by all of the following items except:
 - a. affective reaction.
 - b. schizophrenic reaction.
 - c. schismatic reaction.
 - d. paranoid reaction.
58. Functional psychoses are generally considered to be _____ in origin.
 - a. psychological.
 - b. physiological.
 - c. neurological.
 - d. none of these.
59. An organic agent that is known to cause psychoses is:
 - a. paranoia.
 - b. manic.
 - c. alcoholic.
 - d. schizophrenic.
60. According to the passage, the individual who is characterized by anxiety may be considered:
 - a. psychotic.
 - b. hebephrenic.
 - c. catatonic.
 - d. neurotic.
61. The individual who withdraws and fails to respond to things going on around him may be characterized as:
 - a. depressed.
 - b. psychotic.
 - c. neurotic.
 - d. axiomatic.

(Go on to next page)

62. When one compares psychosis with neurosis, psychosis is:
- a. not as severe.
 - b. nearly as severe.
 - c. more severe.
 - d. about the same severity.
63. It is now known that:
- a. neurosis and psychosis are two different kinds of disorders.
 - b. neurosis and psychosis are the same kind of disorders.
 - c. neurosis and psychosis may be two different kinds of disorders.
 - d. neurosis and psychosis may appear to be in the same person.
64. Psychotic reactions may be classified as:
- a. organic.
 - b. inorganic.
 - c. functional.
 - d. both a and c.

PASSAGE NINE

A photograph of the human brain is more a picture of the cerebral cortex than of anything else, because the cortex incloses almost all the forebrain and midbrain. The cortex, as the picture shows, is like a rumpled piece of cloth that has many ridges and valleys. Anatomists call one of these ridges a gyrus (plural, gyri); a valley or crevice is sometimes called a sulcus (plural, sulci) and sometimes a fissure.

The large sulci or fissures can be used to mark off the cerebral cortex. Along the midline dividing the brain into two symmetrical halves, called cerebral hemispheres, is the longitudinal fissure. Running from this fissure across the top and down the sides of the two hemispheres is the central sulcus. All the cortex in front of this fissure is called the frontal lobe, and this lobe may be considered the expressive part of the brain because it contains motor centers for controlling movements and actions. The cortex behind the central sulcus has been called the receptive part of the cortex, because it contains the centers at which incoming sensory impulses arrive. (There are certain exceptions to these statements, but they serve as a reasonably good way of dividing the functions of the cerebral cortex). Finally, along the side of each hemisphere is a crevice known as the lateral sulcus. The cortex below it and to the side of it makes up the temporal lobe. There are still two other lobes of the cerebral cortex, making four lobes in all, that are not set off by any major fissures. There are the parietal lobe and occipital lobe. The parietal lobe lies immediately behind the central sulcus, and the occipital lobe is the cortex lying under the back of the skull. The functions of these various lobes will be described in the next chapter.

(Go on to next page)

65. This passage is concerned with explaining:
- a. a photograph.
 - b. the cerebral cortex.
 - c. the fissures of the brain.
 - d. the sulci of the brain.
66. There are _____ lobes in the cerebral cortex.
- a. six
 - b. eight
 - c. two
 - d. four
67. The parietal lobe lies:
- a. below the temporal lobe.
 - b. below the occipital lobe.
 - c. in front of the frontal lobe.
 - d. behind the frontal lobe.
68. The expressive part of the brain is found in:
- a. the frontal lobe.
 - b. the parietal lobe.
 - c. the occipital lobe.
 - d. the temporal lobe.
69. The receptive part of the cortex is found in the:
- a. frontal lobe.
 - b. parietal lobe.
 - c. occipital lobe.
 - d. temporal lobe.
70. The cerebral hemispheres are divided by:
- a. the central sulcus.
 - b. the lateral sulcus.
 - c. the longitudinal fissures.
 - d. the major fissure.
71. The name give to the ridges of the cerebral cortex is:
- a. gyri.
 - b. sulcus.
 - c. fissure.
 - d. lobe.
72. The lobe in front of the central sulcus is known as:
- a. the parietal lobe.
 - b. the occipital lobe.
 - c. the frontal lobe.
 - d. the temporal lobe.

(Go on to next page)

PASSAGE TEN

There is one kind of feeble-mindedness which is inherited and is known to be due to a defect in the internal environment. This is phenylpyruvic oligophrenia. It is relatively rare, but its genetic mechanism is reasonably clear. It seems to be caused by a single defective gene which is responsible for an enzyme necessary to utilize phenylpyruvic acid. This acid is a product in the brain's burning of fuel. Ordinarily it is disposed of by a chemical reaction controlled by a specific enzyme. If this reaction is blocked and the acid accumulates in the brain, the result is that the individual is feeble-minded. The diagnostic sign of such feeble-mindedness is excretion of phenylpyruvic acid, for some accumulated acid finds its way to the kidney, and individual who excrete it are without exception feeble-minded.

The internal environment is also clearly involved in another type of feeble-mindedness, cretinism. In this case, the deficiency is in the thyroid hormone. As we have already pointed out the thyroid hormone regulates the rate of metabolism--the utilization of energy--and when it is deficient the brain and other tissues of the body cannot grow or function normally. The child suffering this deficiency becomes a cretin--a feeble minded dwarf with a potbelly and thick, rough skin. Fortunately, we are able to recognize this disease reasonably early and to administer thyroid hormones to alleviate it. Some of the effects of thyroid deficiency on brain function are probably irreversible, especially if the deficiency is severe, but hormone therapy helps the cretin to develop more normally and to have more nearly normal intelligence than he would otherwise. Thyroid deficiency can develop in adults as well as in children. Then it is called myxedema.

73. In feeble-mindedness caused by phenylpyruvic oligophrenia the diagnostic sign is:
- a. large head.
 - b. acid in urine.
 - c. a small head.
 - d. short stubby fingers.
74. Phenylpyruvic oligophrenia is caused by:
- a. heredity.
 - b. indigenous defects.
 - c. a defect in the internal environment.
 - d. a defect in the external environment.
75. Phenylpyruvic oligophrenia is:
- a. common.
 - b. found often.
 - c. relatively common.
 - d. relatively rare.

(Go on to next page)

76. Cretinism is caused by:
 - a. a deficiency in the pituitary hormones.
 - b. a deficiency in the thyroid hormone.
 - c. a deficiency in the adrenal hormones.
 - d. a deficiency in the gonadatropic hormones.
77. A cretin is a feeble-minded:
 - a. giant.
 - b. person of above average height.
 - c. person of average height.
 - d. dwarf.
78. Treatment of the cretin by hormone therapy:
 - a. helps the cretin have nearly normal intelligence.
 - b. helps the cretin have average intelligence.
 - c. helps the cretin have above average intelligence.
 - d. does not help the cretin.
79. The thyroid hormone:
 - a. regulates the rate of metabolism.
 - b. regulates the rate of growth.
 - c. regulates the age at which puberty starts.
 - d. regulates the rate of learning.
80. When a thyroid deficiency develops in adults it is called:
 - a. noxema.
 - b. eczema.
 - c. myxedema.
 - d. cretinism.

PASSAGE ELEVEN

Also of aid in understanding individual persons is the concept of syndrome. This term refers to a pattern of causes and symptoms of disease. In the field of personality it means a pattern of causes and characteristics of the personality. Although each individual's personality is unique, a person may display a syndrome which is similar in many respects to the pattern of characteristics found in other individuals.

Several personality syndromes have been described and measured. In fact, one purpose of some personality tests is to detect such syndromes. For example, a hypochondriacal syndrome, which may consist of many specific characteristics, is one of abnormal concern over bodily health. A psychasthenic syndrome is characterized by excessive doubt, compulsions, obsessions and unreasonable fears. Another syndrome, called the authoritarian personality, is marked by highly conventional behavior, desire for power, hostility, prejudice, and intolerance. This kind of personality pattern appears to have its cause in rejection or excessive domination of a child by his parents.

Our use of the concept of syndrome does not imply, as it might seem to, the classification of people into types, which we criticized earlier in this chapter. Not everyone has a personality syndrome,

(Go on to next page)

and hence syndromes cannot be used to classify people. The idea of the syndrome is that some personality characteristics tend to be highly correlated and thus are similar to those of other individuals. If we are aware of a syndrome, when it exists, we have a better over-all understanding of a person than we would otherwise have.

81. This passage is most concerned with explaining:
 - a. the meaning of personality.
 - b. the classification of personality into types.
 - c. the concept of syndrome.
 - d. the hypochondriacal syndrome.
82. According to the passage:
 - a. everyone has a personality syndrome.
 - b. most everyone has a personality syndrome.
 - c. not everyone has a personality syndrome.
 - d. few have personality syndromes.
83. We have a better understanding of a person when we:
 - a. are aware of a syndrome.
 - b. can classify him into personality types.
 - c. can recognize his uniqueness.
 - d. understand his personality type.
84. The idea of the syndrome is that some personality characteristics tend:
 - a. to be negatively correlated.
 - b. to be highly negatively correlated.
 - c. to be highly correlated.
 - d. to show no correlation.
85. The concept of syndrome refers to:
 - a. similar traits of individuals.
 - b. classification of people into types.
 - c. the etiology of disease.
 - d. a pattern of causes and characteristics.
86. A person may display a syndrome which is _____ the pattern of characteristics found in other individuals.
 - a. exactly like
 - b. nearly like
 - c. completely opposite
 - d. similar to
87. The syndrome characterized by excessive doubt is known as:
 - a. psychasthenic.
 - b. authoritarian.
 - c. Strauss
 - d. hypochondriacal.
88. The syndrome concerned with health is known as:
 - a. psychasthenic.
 - b. authoritarian.
 - c. Strauss.
 - d. hypochondriacal.

(Go on to next page)

PASSAGE TWELVE

Almost every organ of the body has some connection with behavior to the extent that its function is required for the well-being of the individual. Some organs, however, are directly involved in responding to a stimulus. These organs are the sense organs, the nervous system, and the effectors. Sometimes these three parts of the stimulus-response mechanism are referred to respectively as the receptors, adjustors, and effectors. Effectors are of two general types: the muscles and the glands. As indicated elsewhere, the effector glands are of two kinds: the exocrine or duct glands such as the salivary glands, which empty their secretions into cavities of the body, and the endocrine or ductless glands, such as the thyroid gland, which empty their secretions into the blood stream. The endocrine glands are the glands of greatest interest to use because they have widespread effects on behavior.

Not all, though nearly all, of the endocrine glands have nervous connections. Hence not all are effectors. Apart, however, from whether or not they are effectors, their secretions circulating in the blood stream can affect, by stimulation or inhibition, the functioning of other glands. Their secretions can also affect, as we shall see shortly, the activities of many other tissues, including particularly the nervous system and the muscles. They consequently provide a kind of internal environment for the various organs of the body. Internal environment refers to all those chemical, temperature, and stimulus conditions with the body that form an environment for its organs, just as the atmosphere and the external world of stimuli constitute the external environment of the organism.

89. This passage is concerned with explaining:
- organs.
 - the nervous system.
 - behavior.
 - glands.
90. Type (s) of effectors are:
- muscles.
 - glands.
 - neither a nor b.
 - both a and b.
91. The stimulus-response mechanisms are referred to as:
- receptors.
 - effectors.
 - adjustors.
 - all of these.
92. The organs that empty their secretions into cavities of the body are:
- exocrine glands.
 - duct glands.
 - endocrine glands.
 - either a or b.

(Go on to next page)

93. The thyroid gland is:
a. an exocrine gland.
b. a duct gland.
c. an endocrine gland.
d. none of these.
94. Glands of greatest interest because of effects on behavior are:
a. exocrine glands.
b. duct glands.
c. endocrine glands.
d. all of the above.
95. Glands providing an internal environment for the various organs of the body are:
a. exocrine glands.
b. duct glands.
c. endocrine glands.
d. none of these.
96. According to the passage, some organs _____ in behavior:
a. are little involved.
b. are equally involved.
c. more involved.
d. involved completely.

(End of test)

APPENDIX B

COMPLETED TEST OF COMPREHENSION TO
OBTAIN DATA FOR THE STUDY

TEST OF COMPREHENSION

Directions to students: This is a test to determine how well you are able to read selected passages and answer questions that are asked about the passages.

There are twelve reading passages in this test. Read a selection through completely, then answer the questions following it. When you have completed one passage, go to the next. Keep working until you have completed all twelve passages or until you are told to stop. You may look back at the material you have read, if you wish, in order to answer the questions correctly, but do not puzzle too long over any one question. Pass on to the next after a reasonable effort.

Mark all answers on the answer sheet.

PASSAGE ONE

It is often said that man has five senses: sight, hearing, smell, taste, and touch. These five, however, are only the obvious ones we know we have from everyday experience. In addition, there are several more that have been discovered by examining the organs of the body and by experimenting with their function. The fifth of the common five is a skin touch. There are also sense organs in the muscles, tendons, and joints that afford a feeling of pressure within the body. These organs are called kinesthetic receptors, and the sense they serve is kinesthesia. There are also, finally, other sense organs in the head associated with the hearing apparatus. These are called vestibular receptors; they respond to the force of gravity and to rotation of the head. They are the key sense organs in our sense of balance. The real count then is something like ten senses, rather than five. Some of these can be further subdivided into other senses.

The sensitive elements of each sense organ are called receptors. A receptor is a cell that is specialized to respond to relatively small changes of a particular kind of energy. Some receptors, those for sight and smell, for example, are really nerve cells that migrated out from the brain and became specialized for their particular function.

Each of these responds primarily to a particular kind of physical energy. Smell and taste are chemical senses, for they respond to chemical energy. Five senses--touch, pain, kinesthesia, the vestibular sense, and hearing--are mechanical senses, for some kind of mechanical energy is required to activate them. The pain sense, however, may also be stimulated by extremes of chemical and thermal energy. The remaining sense, sight, responds to electromagnetic energy.

1. It is now believed that the number of senses in the human body is:
 - a. five.
 - b. ten.
 - c. something like ten.
 - d. more than ten.
2. Balance is maintained through the:
 - a. vestibular receptors.
 - b. kinesthetic receptors.
 - c. ten senses.
 - d. skin cells.
3. The sense organs in the muscles, tendons, and joints are known as:
 - a. vestibular receptors.
 - b. kinesthetic receptors.
 - c. unspecialized receptors.
 - d. none of these.

(Go on to the next page)

4. Vestibular receptors respond to:
 - a. electromagnetic energy.
 - b. chemical energy.
 - c. force of gravity.
 - d. thermal energy.

PASSAGE TWO

The inner ear is by far the most complicated of the three major parts of the ear. It consists of two kinds of sense organs, one concerned in the sense of balance, and the other in hearing. The organs for balance are called the vestibular sense organs and will be discussed later in this chapter. Because the structure for hearing is spiraled like a snail, it is called the cochlea, meaning snail shell. The cochlea has three different ducts or canals spiraling around together, separated from each other by membranes. Each duct is filled with fluid. Sound vibrations enter them from the ossicles of the middle ear through a membrane known as the oval window, located at the end of the vestibular canal, but too small to show in a drawing. Thus when a sound wave moves the ossicles back and forth, this movement is transmitted to the fluid of the cochlea.

The important event that takes place in the cochlea is the stimulation of nerve cells--called hair cells because they have hairs on their ends--located in the organ of Corti on the basilar membrane separating two of the canals. Pressure changes of the fluid in the canals stimulate these hair cells, causing nerve impulses to be generated and to be conducted through the auditory nerve to the brain.

The basilar membrane beneath the hair cells varies in width from one end of the cochlea to the other, and its width correlates with the frequencies of sounds to which the hair cells are most sensitive. It is narrowest at the end toward the oval window and becomes progressively wider toward the other end.

5. The membrane beneath the hair cells that vary in width from one end of the cochlea to the other is known as:
 - a. the vestibular sense organ.
 - b. the basilar membrane.
 - c. the ossicles membrane.
 - d. the oval window.
6. When a sound wave moves the ossicles back and forth:
 - a. this movement is transmitted to the fluid of the cochlea.
 - b. this movement is transmitted to wax in the middle ear.
 - c. the cochlea stimulates the fluid within it.
 - d. the cochlea is sensitized.
7. The cochlea has three different ducts or canals spiraling around together, separated by:
 - a. bones.
 - b. fluid.
 - c. membranes.
 - d. nerve endings.

(Go on to next page)

8. The most complicated part of the ear is the:
 - a. outer ear.
 - b. inner ear.
 - c. middle ear.
 - d. area between the middle ear and the inner ear.

PASSAGE THREE

Women have a reputation for their ability to talk, and they deserve it. Girls, on the average, learn to talk at an earlier age than boys do. Later, when they can take intelligence tests, they do better on all tests involving the use of language. Thus, they are superior to men in verbal abilities. Closely related to this is the fact that they generally do better on items involving social relations.

Girls are also better on two other related kinds of abilities: perceiving details quickly and accurately and making quick, accurate manual movements. This may account, in part, for the fact that most women are employed in business and industrial work in clerical and secretarial jobs and also for the good record they made during the Second World War in many kinds of industrial jobs. Women also seem to have better immediate memories than men. You will recall that the Stanford-Binet has a series of numbers or words. On the average, women are better in these items than are men.

On the Stanford-Binet test, boys surpass girls on the items that involve spatial, numerical, and mechanical tasks. Correlated with this is the fact that they do better on task involving the perception of spatial relationships, the comprehension of mechanical tools and machines, and mathematical ability. Boys are better in working with numbers and in numerical-reasoning tasks.

We see, then, that men and women differ significantly in certain specific abilities. These conclusions, reached from careful analysis of tests of ability, fit in pretty well with popular conceptions of the difference in ability between men and women.

9. On the Stanford-Binet test, boys surpass girls in:
 - a. items that involve spatial tasks.
 - b. items that involve numerical tasks.
 - c. items that involve mechanical tasks.
 - d. all the above items.
10. Boys do better than girls on tasks that:
 - a. involve spatial relationships.
 - b. involve accurate manual movements.
 - c. involve verbal abilities.
 - d. involve social relations.
11. When girls take intelligence tests, they do better than boys on test involving:
 - a. spatial tasks.
 - b. mathematical ability.
 - c. numerical reasoning.
 - d. the use of language.

(Go on to next page)

12. Girls, being better than boys in perceiving details quickly and accurately and making quick accurate movements, are able to fill successfully such jobs as:
- a. clerical tasks.
 - b. mechanical tasks.
 - c. mathematical tasks.
 - d. spatial tasks.

PASSAGE FOUR

The maze has long been a favorite instrument of psychologists for studying the learning process. A maze is any apparatus consisting of a series of paths and choice points between a starting position and the goal box or outlet. Mazes for animals usually consist of alleys arranged in a pattern of Ys or Ts through which the animal swims, walks, or runs.

As an instrument for studying animal learning, the maze has one unique feature. It provides the opportunity to study serial learning; that is learning a whole series of responses, not just one, leading up to a goal. This raises some interesting questions and throws some light on the processes involved in problem solving. Otherwise, the maze requires instrumental learning just as a Skinner box does, but it also entails perceptual learning. Thus the maze combines serial learning with two varieties of learning we have already studied. By analyzing the processes concerned in maze learning, we gain some understanding of the mechanisms of complex trial-and error learning such as are involved in human thinking and problem solving.

The simplest possible way in which the learning of a maze might take place is that each response at a choice point (a left turn or a right turn) might become associated with the response just preceding it.

13. The unique feature provided by the maze is:
- a. the opportunity to study animal learning.
 - b. the opportunity to watch animals run, swim, or walk.
 - c. the opportunity to study instrumental learning.
 - d. the opportunity to study serial learning.
14. Maze learning involves a response being associated with:
- a. the response just preceding it.
 - b. the two responses preceding it.
 - c. the three responses preceding it.
 - d. none of these.
15. Serial learning is:
- a. learning a response.
 - b. learning a response after the original response.
 - c. learning three responses.
 - d. learning a series of responses.

(Go on to next page)

16. Mazes usually consist of alleys arranged in a pattern of:
- a. Ts
 - b. Ys
 - c. Vs
 - d. a or b

PASSAGE FIVE

Groups of which we are members influence our behavior in various ways, but probably the strongest and most pervasive of the ways are group norms. A norm, as the term implies, is a standard of behavior, but it is more than that. To understand exactly what it is we must refer again to the concept of role and status.

A role is the behavior expected of us in a particular status. The accent should now be placed on the word "expected." A group can expect certain behavior from us because it confers its disapproval on us if we do not do what is expected. Since most of us acquire a need for social approval, and hence do not wish to incur disapproval, we do what is expected of us. In a word, we conform to our group's expectations.

Group norms seem to emerge, like statuses and social structure, wherever a group is formed. A group exists whenever there is any interaction between individuals. By interaction we mean any conversation, exchange of goods and services, or any joint efforts which could tend to cast group members into any kind of status. The longer people interact, and the more they interact, the more they tend to adopt common ways of interpreting the world and the common standards for the behavior of each group member. In other words, group norms to which individuals feel a pressure to conform develop whenever there is any kind of continued interaction among people.

17. Group norms:
- a. are determined by society.
 - b. are determined by group leaders.
 - c. emerge.
 - d. none of these.
18. A group can expect certain behavior from its members because:
- a. it shows disapproval when members disobey rules.
 - b. it sets up norms for the group.
 - c. it rewards the group.
 - d. none of these.
19. Because we wish to be liked by our group:
- a. ~~we conform to our group's expectations.~~
 - b. we interact between individuals.
 - c. we exchange goods or services.
 - d. we propose new types of behavior.
20. A need for social approval for most of us is:
- a. acquired.
 - b. inherited.
 - c. determined by parents.
 - d. determined by teachers.

(Go on to next page)

PASSAGE SIX

From psychological studies of the nervous system of bodily changes in emotion, we know that the changes that occur are initiated by a part of the nervous system called the autonomic system. The changes are therefore called autonomic changes.

The autonomic system consists of many nerves leading from the brain and spinal column out to the various organs of the body, including particularly the blood vessels serving both the interior and exterior muscles. The autonomic system has two parts which usually work in opposition to each other. One part, the sympathetic system, increases the heart rate and blood pressure and distributes blood to the exterior muscles. It is this part that comes into play when we become emotional--or at least when we become fearful or angry. The other part of the system is called the parasympathetic system. It tends to be active when we are calm and relaxed. It does many things that, taken together, build up and conserve the body's stores of energy. For example, it decreases the heart rate, reduces blood pressure, and diverts blood to the digestive tract.

When the sympathetic part of the autonomic system steps up its discharges, as it does in emotion, it produces several symptoms that are worth noting. One set of symptoms concerns the circulation of blood. The blood vessels serving the stomach, intestines, and interior of the body tend to contract in emotion, while those serving the exterior muscles of the trunk and limbs tend to become larger. In this way, blood is diverted from digestive functions to muscular functions to, and the body is thus prepared for action that may involve great muscular activity. At the same time, nervous impulses to the heart make it beat harder and faster, which means that the blood pressure goes up and the pulse rate is quickened. Thus more blood is pumped through the circulatory system to the muscles.

Besides producing changes in circulation, the sympathetic system produces several other bodily changes in emotion. Perhaps, when you have been afraid, you have felt some of them yourself. One is a change of breathing.

21. When the sympathetic part of the autonomic nervous system steps up its discharges, it:
 - a. causes an increase in blood pressure.
 - b. causes more blood to flow to the stomach.
 - c. causes less blood to be pumped to the muscles.
 - d. all of above.
22. A symptom of the sympathetic system is:
 - a. reduced blood pressure.
 - b. decreased heart rate.
 - c. decrease in size of exterior muscles.
 - d. contraction of blood vessels in the stomach.

(Go on to next page)

23. This passage is involved with explaining the functions of the:
- a. central nervous system.
 - b. autonomic nervous system.
 - c. sympathetic nervous system.
 - d. parasympathetic nervous system.
24. The nervous system(s) responsible for increased heart rate is:
- a. the sympathetic system.
 - b. parasympathetic system.
 - c. the central nervous system.
 - d. both a and b.

PASSAGE SEVEN

In order to understand physiological drives we should know the meaning of a concept called homeostasis. This is the tendency of the body to maintain a balance among internal physiological conditions. Such a balance is essential for the individual's survival. Body temperature must not get too high or too low. Blood pressure must not rise or fall beyond certain limits. The blood must not get too acidic or alkaline; it must not contain too much carbon dioxide; it must not become too concentrated; it must have a certain amount of sugar in it. If these limits are exceeded, the individual becomes sick and he may die. Physiologists have discovered that many homeostatic mechanisms are involved in keeping conditions within normal limits. Consider, for example, the control of body temperature. Normal body temperature usually stays near that point because the body can cool or heat itself. If a person's body temperature tends to get too high, he perspires and the resultant evaporation of liquid cools the body. If his temperature threatens to fall, he shivers and steps up his metabolism. Shivering burns the body's fuel faster and thus generates extra heat. In addition, many animals can insulate themselves against heat loss by fluffing their fur and creating a dead-air space around their skin. All that is left of this mechanism in human beings, however, is the goose pimples one gets when he is too cold.

Physiological mechanisms take care of many of the problems of maintaining a homeostatic balance, but the body also makes use of regulatory behavior--behavior that has the effect of regulating internal physiological conditions--to maintain or restore the balance. Such regulatory behavior is behavior that is instrumental in satisfying physiological needs. When the body becomes depleted of water or food, for example, it cannot maintain a balance by calling on its physiological mechanisms. Rather it must obtain more water and food from the outside. It does this through motivated behavior that normally succeeds in procuring more water and food; after that, the homeostatic balance is restored. The important point, then, is that physiological mechanism for maintaining homeostatic balance within the body.

(Go on to next page)

25. Cooling takes place when:
 - a. a body perspires.
 - b. perspiration evaporates.
 - c. a body shivers.
 - d. body fuels burn faster.
26. Body fuels are burned faster when a person:
 - a. shivers.
 - b. perspires.
 - c. rests.
 - d. none of these.
27. Homeostasis may best be explained as:
 - a. control of body temperature.
 - b. the tendency of the body to maintain a balance among internal physiological conditions.
 - c. physiological mechanisms to take care of the problems of depletion of food and water.
 - d. motivated behavior.
28. Regulatory behavior is:
 - a. not motivated by a desire to be comfortable.
 - b. instrumental in satisfying physiological needs.
 - c. instrumental in satisfying psychological needs.
 - d. instrumental in satisfying emotional needs.

PASSAGE EIGHT

For a long time it was thought that neurosis and psychosis were two completely different kinds of disorders. "Once a neurotic, never a psychotic" was almost an axiom. By this, it was meant that neurosis and psychosis were such different reactions that they could not develop in the same person. Today we are not so sure about this. Certainly many patients encountered in clinical practice are difficult to classify and appear to be somewhere in between.

In principle, however, we can make a distinction. Whereas the neurotic individual is characterized by anxiety or strenuous defenses against anxiety, the psychotic individual is typically one who has lost considerable contact with reality. He may simply withdraw and fail to respond to things going on around him. Or he may be so excited or depressed that his reactions are quite inappropriate to circumstances. In many instances, his thought processes may be seriously disturbed by delusions and hallucinations. In any case, he tends to live in a world of his own rather than in the real world around him. For this reason, psychosis is more severe than neurosis, and the psychotic person is more likely to require hospitalization and protective care. Most of the inmates of our mental hospitals, particularly those who stay a long time, are psychotic, not neurotic, individuals.

Psychotic reactions are classified into two categories: functional and organic. The functional psychoses are generally considered psychological in origin, though some of the data on the inheritance of

(Go on to next page)

functional psychoses raise doubts about this. The organic psychoses are those known to be caused by organic agents, such as the syphilis germ, degenerative changes in old age, alcoholism, and the like. As psychologists, we have a greater interest in the functional psychoses. These may be described under three main headings: affective reaction, paranoid reactions, and schizophrenic reactions.

29. It is now known that:
 - a. neurosis and psychosis are two different kinds of disorders.
 - b. neurosis and psychosis are the same kind of disorders.
 - c. neurosis and psychosis may be two different kinds of disorders.
 - d. neurosis and psychosis may appear to be in the same person.
30. An organic agent that is known to cause psychoses is:
 - a. paranoia.
 - b. alcoholic.
 - c. manic.
 - d. schizophrenic.
31. The individual who withdraws and fails to respond to things going on around him may be characterized as:
 - a. depressed.
 - b. neurotic.
 - c. psychotic.
 - d. axiomatic.
32. According to the passage, the individual who is characterized by anxiety may be considered:
 - a. psychotic.
 - b. hebephrenic.
 - c. catatonic.
 - d. neurotic.

PASSAGE NINE

A photograph of the human brain is more a picture of the cerebral cortex than of anything else, because the cortex incloses almost all the forebrain and midbrain. The cortex, as the picture shows, is like a rumpled piece of cloth that has many ridges and valleys. Anatomists call one of these ridges a gyrus (plural, gyri); a valley or crevice is sometimes called a sulcus (plural, sulci) and sometimes a fissure.

The large sulci or fissures can be used to mark off the cerebral cortex. Along the midline dividing the brain into two symmetrical halves, called cerebral hemispheres, is the longitudinal fissure. Running from this fissure across the top and down the sides of the two hemispheres is the central sulcus. All the cortex in front of this fissure is called the frontal lobe, and this lobe may be considered the expressive part of the brain because it contains motor centers for controlling movements and actions. The cortex behind the central sulcus has been called the receptive part of the cortex, because it contains the centers at which incoming sensory impulses arrive. (There are certain exceptions to these statements, but they serve as a reason-

(Go on to next page)

ably good way of dividing the functions of the cerebral cortex). Finally, along the side of each hemisphere is a crevice known as the lateral sulcus. The cortex below it and to the side of it makes up the temporal lobe. There are still two other lobes of the cerebral cortex, making four lobes in all, that are not set off by any major fissures. These are the parietal lobe and occipital lobe. The parietal lobe lies immediately behind the central sulcus, and the occipital lobe is the cortex lying under the back of the skull. The functions of these various lobes will be described in the next chapter.

33. The receptive part of the cortex is found in the:
 - a. frontal lobe.
 - b. parietal lobe.
 - c. occipital lobe.
 - d. temporal lobe.
34. The cerebral hemispheres are divided by:
 - a. the central sulcus.
 - b. the lateral sulcus.
 - c. the longitudinal fissure.
 - d. the major fissure.
35. The name given to the ridges of the cerebral cortex is:
 - a. gyri.
 - b. sulcus.
 - c. fissure.
 - d. lobe.
36. This passage is concerned with explaining:
 - a. a photograph.
 - b. the cerebral cortex.
 - c. the fissures of the brain.
 - d. the sulci of the brain.

PASSAGE TEN

There is one kind of feeble-mindedness which is inherited and is known to be due to a defect in the internal environment. This is phenylpyruvic oligophrenia. It is relatively rare, but its genetic mechanism is reasonably clear. It seems to be caused by a single defective gene which is responsible for an enzyme necessary to utilize phenylpyruvic acid. This acid is a product in the brain's burning of fuel. Ordinarily it is disposed of by a chemical reaction controlled by a specific enzyme. If this reaction is blocked and the acid accumulates in the brain, the result is that the individual is feeble-minded. The diagnostic sign of such feeble-mindedness is excretion of phenylpyruvic acid, for some accumulated acid finds its way to the kidney, and individuals who excrete it are without exception feeble-minded.

The internal environment is also clearly involved in another type of feeble-mindedness, cretinism. In this case, the deficiency is in the thyroid hormone. As we have already pointed out the thyroid hormone regulates the rate of metabolism--the utilization of energy--and

(Go on to next page)

when it is deficient the brain and other tissues of the body cannot grow or function normally. The child suffering this deficiency becomes a cretin--a feeble minded dwarf with a potbelly and thick, rough skin. Fortunately, we are able to recognize this disease reasonably early and to administer thyroid hormones to alleviate it. Some of the effects of thyroid deficiency on brain function are probably irreversible, especially if the deficiency is severe, but hormone therapy helps the cretin to develop more normally and to have more nearly normal intelligence than he would otherwise. Thyroid deficiency can develop in adults as well as in children. Then it is called myxedema.

37. The thyroid hormone:
 - a. regulates the rate of metabolism.
 - b. regulates the rate of growth.
 - c. regulates the age at which puberty starts.
 - d. regulates the rate of learning.
38. Treatment of the cretin by hormone therapy:
 - a. helps the cretin have nearly normal intelligence.
 - b. helps the cretin have average intelligence.
 - c. helps the cretin have above average intelligence.
 - d. does not help the cretin.
39. A cretin is a feeble-minded:
 - a. giant.
 - b. person of above average height.
 - c. person of average height.
 - d. dwarf.
40. In feeble-mindedness caused by phenylpyruvic oligophrenia the diagnostic sign is:
 - a. large head.
 - b. acid in urine.
 - c. a small head.
 - d. short stubby fingers.

PASSAGE ELEVEN

Also of aid in understanding individual persons is the concept of syndrome. This term refers to a pattern of causes and symptoms of disease. In the field of personality it means a pattern of causes and characteristics of the personality. Although each individual's personality is unique, a person may display a syndrome which is similar in many respects to the pattern of characteristics found in other individuals.

Several personality syndromes have been described and measured. In fact, one purpose of some personality tests is to detect such syndromes. For example, a hypochondriacal syndrome, which may consist of many specific characteristics, is one of abnormal concern over bodily health. A psychasthenic syndrome is characterized by excessive doubt, compulsions, obsessions and unreasonable fears. Another syn-

(Go on to next page)

drome, called the authoritarian personality, is marked by highly conventional behavior, desire for power, hostility, prejudice, and intolerance. This kind of personality pattern appears to have its cause in rejection or excessive domination of a child by his parents.

Our use of the concept of syndrome does not imply, as it might seem to, the classification of people into types, which we criticized earlier in this chapter. Not everyone has a personality syndrome, and hence syndromes cannot be used to classify people. The idea of the syndrome is that some personality characteristics tend to be highly correlated and thus are similar to those of other individuals. If we are aware of a syndrome, when it exists, we have a better over-all understanding of a person than we would otherwise have.

41. The concept of syndrome refers to:
 - a. similar traits of individuals.
 - b. classification of people into types.
 - c. the etiology of disease.
 - d. a pattern of causes and characteristics.
42. We have a better understanding of a person when we:
 - a. are aware of a syndrome.
 - b. can classify him into personality types.
 - c. can recognize his uniqueness.
 - d. understand his personality type.
43. The syndrome concerned with health is known as:
 - a. psychasthenic.
 - b. authoritarian.
 - c. Strauss.
 - d. hypochondriacal.
44. The syndrome characterized by excessive doubt is known as:
 - a. psychasthenic.
 - b. authoritarian.
 - c. Strauss.
 - d. hypochondriacal.

PASSAGE TWELVE

Almost every organ of the body has some connection with behavior to the extent that its function is required for the well-being of the individual. Some organs, however, are directly involved in responding to a stimulus. These organs are the sense organs, the nervous system, and the effectors. Sometimes these three parts of the stimulus-response mechanism are referred to respectively as the receptors, adjustors, and effectors. Effectors are of two general types: the muscles and the glands. As indicated elsewhere, the effector glands are of two kinds: the exocrine or duct glands such as the salivary glands, which empty their secretions into cavities of the body, and the endocrine or ductless glands, such as the thyroid gland, which empty their secretions into the blood stream. The endocrine glands are the glands of greatest interest to us because they have widespread effects on behavior.

Not all, though nearly all, of the endocrine glands have nervous connections. Hence not all are effectors. Apart, however, from

(Go on to next page)

whether or not they are effectors, their secretions circulating in the blood stream can affect, by stimulation or inhibition, the functioning of other glands. Their secretions can also affect, as we shall see shortly, the activities of many other tissues, including particularly the nervous system and the muscles. They consequently provide a kind of internal environment for the various organs of the body. Internal environment refers to all those chemical, temperature, and stimulus conditions with the body that form an environment for its organs, just as the atmosphere and the external world of stimuli constitute the external environment of the organism.

45. According to the passage, some organs _____ in behavior.
- a. are little involved
 - b. are equally involved
 - c. are more involved
 - d. involved completely
46. The stimulus-response mechanisms are referred to as:
- a. receptors.
 - b. effectors.
 - c. adjustors.
 - d. all of these.
47. Type(s) of effectors are:
- a. muscles.
 - b. glands.
 - c. neither a nor b.
 - d. both a and b.
48. The thyroid gland is:
- a. an exocrine gland.
 - b. a duct gland.
 - c. an endocrine gland.
 - d. none of these.

(End of test)

VITA

WAYNE DALE LEE

Candidate for the Degree of

Doctor of Education

Dissertation: READABILITY OF GENERAL PSYCHOLOGY TEXTBOOK MATERIAL:
A CROSS-VALIDATION STUDY OF THE DALE-CHALL READABILITY
FORMULA

Major Field: Higher Education - Reading

Biographical:

Personal Data: Born at Siloam Springs, Arkansas, December 16,
1926, the son of William Sidonal and Lona Mae Anglin Lee.

Education: Attended grade schools in Arizona, California, and
Oklahoma; graduate from Crowder High School, Crowder,
Oklahoma, in 1945; received Bachelor of Science in Educa-
tion degree from Southeastern State College, Durant, Okla-
homa, with a major in mathematics and chemistry in May,
1955; received the Master of Teaching degree from South-
eastern State College, Durant, Oklahoma, with a major in
Elementary Education in July, 1958; completed requirements
for the Doctor of Education degree in May, 1963.

Professional Experience: Entered the Army of the United States
in 1945 and was discharged in 1946 with the rank of ser-
geant; taught grades 1 to 8 at Taloga Elementary School,
Richfield, Kansas, 1950-1951; taught mathematics, science,
and coached at Nida High School, Kenefic, Oklahoma, 1951-
1952; taught grades 5 to 8 at Pleasant Valley Elementary
School, Ulysses, Kansas, 1953-1954; taught 6th grade and
supervised playground at George Washington Elementary
School, Durant, Oklahoma, 1954-1959; served as graduate-
assistant at the Oklahoma State University, 1959-1961;
served as Instructor of Elementary Education at Central
State College, Edmond, Oklahoma, 1961-1963.

Member of Phi Delta Kappa, Life Member of National Education
Association, Oklahoma Education Association, and Assoc-
iation of Student Teachers.