THE DEVELOPMENT OF AN INSTRUMENT TO DETERMINE THE

STUDY SKILL OF COLLEGE FRESHMEN

APPROVED:

Graduate Committee:

[Signatures]

Major Professor

[Signature]

Committee Member

[Signature]

Committee Member

Dean of the School of Education

[Signature]

Dean of the Graduate School
THE DEVELOPMENT OF AN INSTRUMENT TO DETERMINE THE

STUDY SKILL OF COLLEGE FRESHMEN

DISSERTATION

Presented to the Graduate Council of the
North Texas State University in Partial
Fulfillment of the Requirements

For the Degree of

DOCTOR OF EDUCATION

By

John D. Polk, B. A.

Denton, Texas

May, 1965
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
</tbody>
</table>
| Definition of Terms
Statement of the Problem
Hypotheses          |    |
| II. SURVEY OF LITERATURE | 8  |
| Reading
Habits and Attitudes
Organization | |
| III. METHOD OF DEVELOPMENT OF STUDY SKILLS TEST | 31 |
| Item Selection
Pilot Study
Statistical Analysis
Reliability
Validity
Other Criteria
Subjects
Collection of Data | |
| IV. RESULTS | 67 |
| V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS | 76 |
| APPENDIX | 84 |
| BIBLIOGRAPHY | 98 |

iii
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Intercorrelation Matrix of Variables for 120 Subjects.</td>
<td>49</td>
</tr>
<tr>
<td>II. Intercorrelation Matrix of Variables for 62 Males.</td>
<td>51</td>
</tr>
<tr>
<td>III. Intercorrelation Matrix of Variables for 58 Females.</td>
<td>52</td>
</tr>
<tr>
<td>IV. Actual and Predicted Relationships of Study Skills and Predictor Variables.</td>
<td>69</td>
</tr>
<tr>
<td>V. The F Level of Significance for Six Predictor Variables for the Sample of 120 Subjects.</td>
<td>70</td>
</tr>
<tr>
<td>VI. The F Level of Significance for Six Predictor Variables for the Sample of 58 Female Subjects.</td>
<td>71</td>
</tr>
<tr>
<td>VII. The F Level of Significance for Six Predictor Variables for the Sample of 62 Male Subjects.</td>
<td>72</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Study skill is essential to success in college. The ability to organize, evaluate, interpret and draw conclusions from printed material is the skill that the successful student must possess.

Study skill has been defined as reading skill, mental ability, critical thinking, habits, and attitudes. Tests in each of these areas have been developed, with some statistical evidence that a measure of one or more of these concepts is related to academic success.

The development of study skill tests have grown out of the increasing interest in the prediction of academic success resulting from study habits and attitudes of college students. Entwisle's (20) study of twenty-two evaluation studies shows that there is a widespread interest in study habits. Blake's (3) 1953 study found about 90 per cent of the American colleges were interested in study habits. Bennett's (1) interest led to a study in which she found that students have more difficulty in the area
of study skill than any other area of college life. Just as one would expect, this interest has not brought all writers to the same conclusion.

Reading as a study skill is of primary value to most writers interested in academic prediction. Reading skill is basic to all other skills developed by the college freshman. Almost as important as reading skill is the effective planning and organization of material. Book (4), Bennett (1), Chandler (5), Bird (2), and Jordan (8) say that organization is the foundation of learning.

Chandler and others (5) seem to have the most inclusive list of study skills. They include reading, writing, locating information, note taking, listening, preparing lesson assignments, recall of information and facts, taking examinations, using the literature, preparing written reports, participation in discussion, and making oral reports. It would seem, then, that study skill is all of the things a student can utilize in the mastery of academic material. These skills are so general that they are difficult to determine.

Several tests have been designed to measure both habits and skills. Holtzman and Brown, California Test Bureau, Tyler-Kimber and Wrenn have all developed tests
in this area. One of the more recent study-skills tests to be developed is a sub-test of the Contemporary Tests of Scholastic Progress for the seventh-grade level published by the California Test Bureau in 1962. The administrator's Manual (6) states that 6,480 students were used in arriving at the norms for the test. In the author's judgment this is the most adequate study skills test that has been developed to date. There is, however, a need for a study skills test for college students.

A college student should be able to organize printed material into outlines by reading ideas rather than words and phrases. A study skills test should be useful in measuring the student's ability to choose the central idea and the related areas from a paragraph. He should be able to evaluate the material.

Evaluation of material involves the choice of relevant material and exclusion of that which is not relevant to the central thought of the topic being considered. Interpretation is related to both organization and evaluation.

Interpretation is essential to skill in understanding printed material and involves the use of material in a meaningful way. The ability to interpret printed material
is essential to all academic achievement. Understanding is basic to drawing conclusions.

The items in the author's study skills test used to determine the student's ability to organize are items based on the correct choice of an outline made after reading a paragraph. The student's ability to evaluate is based on the correct choice of answer on items that have one central thought. Ability to interpret is based on the correct choice of answers on items concerned with the meaning of a paragraph. Ability to draw conclusions is based on items where graphs are used. There are various items designed to measure each of the above concepts.

A study skills test of this kind is needed to determine the study skill of college students. It will be useful in screening those students who have average ability but are unable to succeed in college because of poor study skill.

Definition of Terms

1. **Scholastic achievement.**—The ability to understand ideas expressed in words, the ability to work with numbers, the ability to solve logical problems, and the ability to visualize objects and figures in space as these concepts are measured by the PMA test.
2. Critical thinking. -- The ability to reason as it is determined by the Watson-Glaser Critical Thinking Appraisal.

3. School grades. -- The average of each student's grades for one semester. Physical Education is not included.

4. Study skill. -- The ability to organize, evaluate, interpret and draw conclusions from printed material as it is measured by the Study Skill Test.

Statement of the Problem

The purpose of this study was to develop an instrument to determine the study skills of college freshmen. The following sub-problems were involved:

1. To determine the reliability of the test for internal consistency obtained throughout the test and the consistency with which the test measures from one administration to another.

2. To determine the internal validity by item analysis and predictive validity by correlation with other academic-success predictors.
Hypotheses

The following hypotheses were formulated and investigated by statistical analysis of the data collected.

1. The instrument developed for this study will be a reliable study skill test.

2. The relationship between study skill as measured by the test and the six predictors of academic success listed below will be positive and in the following rank order of magnitude from highest to lowest.
   a. School grades
   b. Critical thinking scores
   c. PMA Verbal meaning
   d. PMA Reasoning
   e. PMA Spatial relations
   f. PMA Number facility

3. The combination of variables most significantly related to study skill will be school grades, critical thinking, and verbal meaning.
CHAPTER BIBLIOGRAPHY


CHAPTER II

SURVEY OF LITERATURE

Although the writers cannot be divided into distinct categories by the criteria they use in the prediction of academic success, they do follow some rather general lines of preference. Even though there is some overlapping among the criteria, these can be divided in a limited way. The variables used most to predict college success are reading ability, habits and attitudes, organizational skill and critical thinking which is a more recent variable chosen as a predictor of academic success. Many of the "How-to-Study" courses taught in college include at least one chapter on the value of reading ability.

Reading

Robinson (56) states that some college freshmen can read no better than the average fourth-grader. Studies show that scores on tests of reading ability may be as highly related to college success as are the scores on tests of intelligence. Robinson (55) includes intellectual ability, slow reading rate, small vocabulary, and
poor organizing skill as being determining factors of reading ability. About one fourth of his book is used for teaching reading skill. Bennett (5) indicates that Indiana University offered courses in "How to Study" which helped some students increase reading ability. Some of the gains in reading speed were as much as 250 per cent. Skill in reading is an important tool of the college student since every field of study involves the securing of facts from reading. He concluded that:

... the ability to read rapidly and effectively has much to do with a student's success in college, and many freshman students do not possess sufficient skill in the art of reading to enable them to succeed with their academic work (5, p. 165).

Bennett cites another study that indicates the variation between students who succeed and those who do not is greater in reading ability than in any other factor. More recent studies have verified these results.

In an attempt to determine academic achievement at the college level, Endler (2, p. 695) found that the best predictor of freshmen grade-point averages to be high school grade-point averages and the next best predictor was a reading score from the Sequential Tests of Educational Progress. However, McQueen (4, p. 629) reports finding test scores and grade-point averages for 246 University of Nevada freshmen.
There are few studies that have agreed with McQueen's findings. Many studies have been made to determine the relationship between remedial reading course scores and academic success.

Instructions in higher level reading skills have been given by several colleges. Robinson (5, p. 29) and Kilby (10) insist that such instruction given at the high school level would help one predict more successful achievement in college. Remedial reading course scores and grade-point averages were significantly related and experimental groups showed an improvement in grade-point averages significant at the .05 level in two separate studies made by Mouly (6) and Barbe (7). Since Mouly indicated that improvement was found only for those "who took the course seriously," he felt that there could be some variable involved in the improved grade-point average other than just reading skill. However, McGinnis (47) is more positive in her evaluation of remedial reading, indicating that the entire experimental group made higher point-hour-ratios than the control group. A few studies define the area of reading skill which must be improved to increase the grade-point averages.

Improvement in reading speed alone is not enough. Comprehension is essential to an improvement in grade-point
averages that is a result of reading. Kingston and George (40) insisted that only a curriculum that is largely linguistic in nature would benefit from a college reading program. Comprehension, vocabulary, and rate of reading are all related to academic success. Ransom (54) stresses rate of reading less than the other two factors. Blake (12) rates comprehension as the most important aspect of reading skill.

Sparks (59) in a study of high school reading ability found that skills involved were: vocabulary, concentration, organization, spelling, and summarizing as well as reading rate and comprehension. Kingston and George (40) found this same thing in a study of engineering students. Seegars (58) found high level verbal understanding a more accurate predictor of school grades than an above average intelligence score. Laroe (42) in a study of school children's study habits had the children make a list of what they thought would be proper study habits. Of the fifty-eight items in the list only eight were related to reading. Much more attention was given to habits and attitudes than to reading ability.
Habits and Attitudes

Habits and attitudes have received wide attention in an effort to find accurate predictors of academic success. Josephenia (37) uses study habits and study skills interchangeably in pointing out that these habits or skills are the vehicles to academic success. These habits include the ability to apply the tools of research, the needed discipline to gather, synthesize, organize, and present data. Xavier (67) indicates that she sees student's attitude toward study as more important than any specific study habit. Although Michael and Reeder (43) found a zero correlation between a study habits inventory and verbal skill, they found the correlation to be significant between study habits and grade averages. They indicate that study habits scores could profitably be used in predicting grade averages. Brown and Dubois (10) support the theory that study habits can be good predictors of academic success but they include motivation and experience as additional factors that must be present. They admit that the experience might well be a change in the perception of study. Other studies have found the relationship between study habits and attitude inventory scores and predictors of school grades such as achievement tests rather low.
In support of study habits and attitudes as predictor variables for academic success, Holtzman and Brown (34) point out the unreliability of one-semester grades. When they used the American College Examination and Survey of Study Habits and Attitudes scores in multiple correlation, they found that the Survey of Study Habits and Attitudes scores added significantly to the result. The multiple correlation was reported as .63 for females and .73 for males. There has been considerable investigation in an effort to find non-intellectual correlates of academic success. Brown (9) believes the most promising in this area both logically and empirically has been study habits, even though his study at Iowa State University showed that study habits could not be used to predict grades. Chahbazi (21) and Ahman (1) found about the same result. From these findings, there arose an interest in other non-intellectual predictors of school grades.

Attitudes was one of the centers of interest for those who found habit alone insufficient for predicting academic success. Brown and Holtzman (11) investigated the Study Habits Inventory of Wrenn and found little relationship between those scores and academic success. They found that
an attitude index could be designed that would correlate favorably. At the University of Texas, a correlation of .46 was found between attitudes and grade-point averages. At San Antonio College there was a correlation of .56 between attitudes and grades. It was an easy transition from habits and attitudes to other factors in the prediction of grades.

Organization

The organization of study has been investigated by several writers. Carter (18) found that a self-report inventory of organization of study procedure correlated .41 with school achievement. Muse (46), an early investigator of study habit, found that grades were related to the number of times a student read an assignment. Tussing (63) insists that behavior patterns developed in study is the way to academic success. Behavioral patterns are the actual mechanics of study. Study procedure test scores are more effective predictors of academic achievement than intelligence test scores. Carter (17) found a correlation of .32 between study methods test scores and grade-point averages. He defines methods as planning, thought, and mechanical aspects of study. Stokes (61) includes meditation,
reflection, and self-improvement in his list of methods.

McLane (48) has a longer list. He includes: correct study habits, develop a desire to learn, pay attention in class, realize the importance of home work, and prepare adequately for tests. Barton (3), Beamer (4), DiMichael (27), Flemming (32), Cheyney (23), Mulholland (45), Johnson (35), Jordan (36), and Charles (22), include note taking and outlining in their lists of study techniques. Reading skill, study habits and attitudes, and methods or skills have all been included in one study improvement course or another. A review of several study skills courses is enlightening.

Entwisle (31) reviewed twenty-two study skills courses and found that grade-point average was the criterion measure in all but two of the study skills courses. She further found that in every study reviewed there had been some improvement in the students' grades. There was not a great deal of difference between one type of study skill and another in the effects on students' grade-point averages. The criterion for success of a study skills course was the student's desire to take the course. The value of several different measures to determine study skill has led to the development of a large number of such tests.
The following tests are listed by Buros (13, 14, 15, 16) as study skills tests: Study Outline Test, 1926; Study Habits Inventory, 1934; The Tyler-Kimber Study Skills Test, 1937; the Traxler High School Reading Test, 1939; Test of Study Skills, 1940; Work Study Skills, 1940; The Use of Library and Study Materials, 1941; Study Habits Inventory, Revised Edition, 1941; The Watson-Glasser Test of Critical Thinking, 1942; Survey of Study Habits, Experimental Edition, 1944; Work Study Skills, 1947; Bennett Use of Library Test, 1947; Cooperative Dictionary Test, 1951; Edminster How to Study Test, 1947; Interpretation of Data Test, 1950; Logical Reasoning Test, 1950; Test on the Use of Books and Libraries, 1950; Test of Critical Thinking, 1951; Uses of Sources of Information, 1951; Spitzer Study Skills Test: Evaluation and Adjustment Series, 1955; Brown-Holtzman Survey of Study Habits and Attitudes, 1956; SRA Achievement Series: Work Study Skills, 1957; California Study Methods Survey, 1958; Watson-Glasser Critical Thinking Appraisal, 1956; and a long list of others that are sub-tests of achievement test series. These are a few of the tests that seem to be representative of the list of tests.

One of the more popular reading tests is the Nelson-Denny Reading Test (51). Booker's review in Buros' (16)
work indicates that it has been outstanding for more than twenty years. Although the Nelson-Denny Reading Test was designed for use at all educational levels, it has proven more effective at the college freshman level. This two-part test measures both vocabulary and comprehension. The vocabulary section of 100 test items is not designed to measure speed or comprehension. The second part of the test is designed to measure only comprehension. Of the 136 test items only thirty-six are designed to measure comprehension, while there are 100 items in the vocabulary section. There are nine paragraphs of about 200 words each with four test items for each paragraph. The reliability coefficient of the test determined by the test-retest method is .90. The validity of the test rests on the fact that it serves the purpose for which it was designed. There are other reading tests that give several measures.

The SRA Reading Record (57) for grades 8-13 was developed in 1947 to give eleven different measures related to reading ability. They were: reading speed, comprehension, paragraph meaning, directory reading, map-table-graph reading, advertisement reading, general vocabulary, sentence meaning, technical vocabulary, and total score. Turnbull's review in Buros' (14) work calls these the ten basic skills
of reading. He indicates that this is really little more
than a basic reading test. The test of vocabulary skill
is the most adequate of the eleven tests. The reliability
coefficient is reported to be high and the test measures
what the designer says it will measure. Since the pub-
lisher of the test has developed several aids to reading
improvement and uses the test as a measure of the need for
remedial work, it seems to be an acceptable instrument that
is rather representative in its field. For those who feel
that reading ability is the essential nature of study
skills, these two tests would be good measures for the
prediction of school grades. Both the Nelson-Denny Reading
Test and The SRA Reading Record are recognized as adequate
measures of the skill connected with reading ability.
Buros (14) has reviews that recommend these tests. These
tests would be inadequate for those who predict school
grades on the basis of habits and attitudes.

The Study Habits Inventory (62) developed by Wrenn
and revised in 1941 has been rather popular in this field.
Jones (15) describes the test as divided into four sections,
with each section covering a different set of habits. Sec-
tion "A" is Reading and Note Taking Technique. It is that
section that relates it to the reading test field. Section "B" is Habits of Concentration. Section "C" is Distribution of Time and Social Relationships in Study. Section "D" is General Habits and Attitudes of Work. Section "D" relates this test to the habits and attitude tests. There are twenty-eight weighted items on the inventory which is a self-report. Most reviewers in Buros (14) say that self-reports are the most difficult to test for validity. A self-report assumes honesty and intelligence on the part of the subjects. One has no guarantee of either in any subject. One reviewer (14) has said subjects will answer a self-report to please the tester, if he likes the tester, or he will answer them so he will be enhanced. Either way, such answers could not lead to the validity of a self-report as a measuring instrument. Even Brown and Holtzman (11) indicate that the Study Habits Inventory is a poor measuring device. They cite some studies that have found little or no relationship between the Study Habits Inventory and scholastic success. They developed a study habits instrument themselves in an effort to improve prediction.

The Brown-Holtzman Survey of Study Habits and Attitudes (12) is probably the most popular instrument of its kind.
Known as the SSHA, it has been the tool in many research projects in education. It is a self-report questionnaire of seventy-five items designed to help the student identify his habits and attitudes whether they are bad or good. In studies previously reported, Brown (11) reports several instances of the correlation between SSHA and grade-point averages that are .50 and above. He reports a somewhat higher correlation coefficient for females than for males. In another study reported by Holtzman and Brown (34) they indicate that .40 male subjects taking the American College Examination and the SSHA had a multiple correlation of .73. In a combination of variables the Survey of Study Habits and Attitudes would appear to make a significant contribution to the prediction of school success. However, when it is used by itself it has little or no predictive value. This was verified by both Brown (11) and Ahman and Glock (1). If study skill is to be used for predicting academic success it must be more than reading ability and habits and attitudes.

Some of the test designers have come out with more comprehensive instruments. A large number of study skills tests have been designed. Only a few will be considered here. One of the first study skills tests was developed
by Tyler and Kimber. The Tyler-Kimber Study Skills Test (64) was copyrighted in 1937. It includes the eight following sub-tests:

2. Using an index.
4. Recognizing common abbreviations.
5. Using the library card catalog.
6. Interpreting maps.
7. Knowing current periodical literature.
8. Interpreting graphs.

It may be seen from the areas included in the sub-tests of the Tyler-Kimber Study Skills Test that it is not a measure of reading ability nor a habits and attitudes inventory. It is a test of the subject's ability to handle the tools that are essential to academic success of the college student. The area covered by each of the sub-tests is an essential part of the skill used in school work. No one course in college would require the use of all the skills included in the test. There is no course in college that could be mastered without some of them.

The Spitzer Study Skills Test is a more recent development. It was copyrighted in 1956. This test is designed
to measure the ability of students to use four important work-study skills. They are: using a dictionary; using an index; locating sources of information; understanding graphs, tables, and maps; and organization of facts and note taking. Constructed for use with high school students and college freshmen, the test provides an over-all index of achievement as well as a reasonably accurate measure of each of the listed skills. There are about 170 items in the test. The reliability coefficient determined by the split-half method is about .90. Since each of the sub-tests measure a different concept, each of them has a validity coefficient ranging from .00 to .70. A more recent study skills test has been developed by the California Test Bureau.

Contemporary Tests of Scholastic Progress (25) published in 1962 includes a study skills test containing items from the following areas: items based on maps, graphs, tables, charts; vocabulary reference skills; use of index; use of card catalog; and dictionary skills. This test was experimentally administered to nineteen different Texas schools. The scored tests were subjected to an item analysis and statistics were computed for each item. The per cent correct response was determined by computing the
per cent of students who answered the item correctly. The per cent of discrimination was determined by subtracting the bottom 27 per cent of the students from the top 27 per cent in terms of correct response. When the tests were evaluated by a group of raters, 88 per cent agreed the test was of major importance. This significant test measures the study skill of elementary and secondary school pupils. All of the above tests are valuable in the area each of them is designed to measure.

In the measurement of reading skill, the Nelson-Denny Reading Test (51) has been very effective for the last twenty years (16). It was designed to measure speed and comprehension in reading not as a study skills test. Along the same line of measurement the SRA Reading Record has been very effective (14). Although it is more inclusive in the areas tested, it is a reading test rather than a study skills test. Wrenn's test, The Study Habits Inventory (15), does not include all of the skills essential to academic success. The Survey of Study Habits and Attitudes (12) is a self-rating instrument and does not measure study skill. Both the Tyler-Kimber Study Skills Test (64) and the Spitzer Study Skills Test (60) are measures of study skill. The Contemporary Tests of Scholastic Progress (25)
is the most recent addition to the field. This test is for use at the high school level. There is a need for a study skills test at the college level.

The purpose of this study was to develop an instrument for use with college freshmen. Although some of the instruments cited in this study are both valid and reliable, they do not meet the need for a study skills test at the college level. In addition to items measuring verbal skill, reference skill, indexes and card catalogs used in other study skills tests, this text includes items requiring thinking and reasoning. This test was designed to be more discriminating than most of the existing study skills tests.
CHAPTER BIBLIOGRAPHY


2. Barbe, W. B., "The Effectiveness of Work in Remedial Reading at the College Level," *Journal of Educational Psychology*, XLIII (July, 1952), 229-237.


42. Laroe, Willetta R., "How to Study," *Grade Teacher*, LXXI (October, 1953), 56.


50. McQueen, R., "Diagnostic Reading Scores and College Achievement," *Psychological Reports*, III (May, 1957), 627-629.


CHAPTER III

METHOD OF DEVELOPMENT OF STUDY SKILLS TEST

The study of the strengths and weaknesses of the existing test instruments reveals the need for an improvement in the design of an instrument to measure the study skill of college freshmen.

The instrument to measure study skill was developed by first planning the test to meet the need for predicting academic success for college freshmen. The data were gathered by using samples of subjects from the population to which the test would be applicable. The subjects were 257 college freshmen who provided the test scores for determining the reliability and validity coefficients. These coefficients were determined by statistical analysis of the data.

Item Selection

One reason for the development of the study skills test was the need for an instrument to better predict academic success for college freshmen. This problem has
probably received more attention than any other single problem in education (6, p. 6).

The development of the study skills test involved at least three distinct steps: planning, construction, and evaluation. According to Green (12, p. 13), a good test does not happen by chance, but is the result of extensive planning, thought, and work.

Several of Green's (12) general principles of test construction were followed in planning the study skills test. The test items were planned in such a way that reading rate and comprehension would not unduly influence the test scores since the test was not designed to measure these variables. Only multiple-choice items were planned with a simple method of indicating responses. They were to have four possible responses for each test item. Wood (23, p. 26) states that one of the basic characteristics of a multiple-choice item is that the item itself contains the standard by means of which the best answer is to be chosen. The problem can be posed either by direct question or by incomplete statement. Both types were planned for the study skills test. Each item was planned to be appropriate to the ability of college freshmen. The transition
from planning of the test to construction is simplified.

Bean says:

Having the tentative outline before him the research worker in test development is now ready to obtain items from whatever source he can discover, buy, borrow, steal, or invent. Whatever his tactics may be, he will have the responsibility of evaluating what he compiles as he goes along (3, p. 37).

The test items came, primarily, from two sources--the author, and a large number of college students and experienced teachers. No one individual developed enough items so that the total test reflected individual influence.

The study skills test was developed from items prepared by graduate students at the North Texas State University enrolled in counseling and guidance. These students were instructed to write test items which would measure the study skill of college freshmen. The items were prepared over a period of two years by several different groups of students and experienced teachers.

Each group submitted the items to the author at the end of each school semester. All items were then typed on eight and one half by eleven inch sheets of paper. Four items were written on each sheet of paper. The sheets were divided into four pieces with one test item on each piece. All items were then reviewed carefully by a jury of graduate
students and all items that had more than one correct answer as well as the items that had no correct answers were discarded. All items that were obviously unrelated to the study skills test were discarded.

After all unusable items were removed from the total number of available ones, there were about 500 left from which to choose the final items used in the test. These items were organized into related groups by a jury of thirty-five graduate students and experienced teachers.

To facilitate the organization of the test items into related groups, a box was divided into seven compartments. There was one compartment for each of the categories listed in Bloom's (5) Educational Objectives and one compartment for the items that could not be categorized. The educational objectives are knowledge, comprehension, application, analysis, synthesis, and evaluation.

Bloom (5) has subdivided the list as follows:

1. Knowledge
   a. Knowledge of specifics
   b. Knowledge of ways and means
   c. Knowledge of universals and abstractions

2. Comprehension
3. Application

4. Analysis
   a. Analysis of elements
   b. Analysis of relationships
   c. Analysis of organizational principles

5. Synthesis
   a. Production of a unique communication
   b. Production of a plan
   c. Production of a set of abstract relationships

6. Evaluation
   a. Judgments in terms of internal evidence
   b. Judgments in terms of external evidence

The subdivisions were used as guides in item selection. However, only the major divisions were used as headings for the compartments of the box used for the organization of the items.

Each item was placed in the compartment which best described the type of information needed to answer it correctly.
The first draft of the test was made by selecting the twenty items that were most like the heading from each of the six major divisions of Bloom's (5) Educational Objectives. That it might be limited to chance that two items from the same division appeared together on the completed test, the 120 items were placed in a large box and thoroughly mixed together. The items were then taken from the box one at a time and placed in the order in which they were drawn in the first draft of the test. The first draft of the test was a mimeographed form of 120 questions which had four possible answers for each question. Only one of the possible answers would answer each question correctly.

Pilot Study

The first draft of 120 items was used in a pilot study made at the North Texas State University in December, 1963.

The pilot study is used usually with only a small number of subjects, to suggest what specific values should be assigned to the variables being studied, to try out certain procedures to see how well they work and more generally to find out what mistakes might be made in conducting the actual experiment so that the experimenter can be ready for them (15, p. 49). In the pilot study none of the
subjects completed the 120 items on either the test or retest phase of the study. From this study it was assumed that there were too many items for even the superior subject to respond to in the fifty minutes allowed for the average freshman class period. Since there were not enough scores completed on items numbering above 100 for an estimate of reliability, the entire twenty items were removed from the test. Three of these items were used later to replace items that were unacceptable.

The reliability of the test in this pilot study was determined by the test-retest method. The correlation between the test and retest scores was .7322. The mean for the test was 38.23 and the mean for the retest was 49.06.

The validity of the test was determined by using the Watson-Glaser Critical Thinking Appraisal as a criterion. The correlation between the study skill scores and the criterion scores was .5294. Cronbach (18, p. 132) states, "The correlation between the test and an independent criterion can never be higher than the square root of the correlation between two forms of the test." The square root of the correlation between the two forms of the test was .5361, and therefore comes within the limit of this rule.
The results of the pilot study showed that twenty of the items were unacceptable for the test. Fourteen of the twenty items were answered correctly by the entire group of subjects. Six items were answered correctly by students whose scores were below average for the group and were assumed to be chance responses. The removal of the twenty items left eighty items for the test to be used with the first group of subjects to take part in the reliability and validity studies.

The scores of the group of 100 subjects on the eighty-item test were divided into two separate groups. The mean for this total group was determined and the scores falling on the mean or above were placed in one group and those falling below the mean were placed in another group. Two lists were made with numbers corresponding to the test items numbers. On one of the lists the correct responses of the above-average group were tallied and on the other list the correct responses of the below-average group were tallied.

The tallies revealed that ten items were responded to incorrectly by 80 per cent of the above-average group and twenty items were responded to correctly by 75 per cent of the below-average group. These thirty items were removed
from the test leaving only fifty of the original 120 items. This test was then used with the subjects chosen for the study of validity and reliability.

Green (12, p. 17) says that a test should be long enough to be valid and reliable but short enough to be usable. There is a point of diminishing returns where lengthening the test will lower reliability since fatigue of the subjects affects test results. The length of the test must be a compromise between the considerations of validity, reliability, and usability.

Statistical Analysis

The pilot study of the test conducted at the North Texas State University gave some worthwhile information. Thirty subjects were given the test and one week later were retested. This statistical procedure determined the equivalence coefficient of the test (4). The reliability coefficient was .7322. The preliminary study of validity was also made in the pilot study.

Study skill is in part defined as the ability to think. Since this ability is included in the assumptions of the Watson-Glaser Critical Thinking Appraisal, it was used as a criterion in the study of validity. A validity
coefficient of .5294 was determined by a correlation be-
tween study skills and the Critical Thinking Appraisal. 
Both the coefficient of reliability and the coefficient 
of validity are significant. With 28 degrees of freedom 
(n-2), Garrett (11, p. 201) says that even a correlation 
of .4500 would be significant at the .01 level. Both the 
reliability and validity coefficient exceed this level. 
The study was continued.

According to Stevens (4, p. 1243) reliability is the 
consistency or stability of the evaluation obtained from 
repeated observations. Estimates of the stability classi-
fications over a sample of similar tasks are provided by 
two coefficients, equivalence and stability.

Reliability

The tenability of hypothesis one of this study was 
determined by examining the data and treating it statisti-
cally for reliability. The coefficient of reliability was 
determined.

1. The coefficient of reliability was obtained by 
the test-retest method.

2. The coefficient of reliability was further veri-
fied by the split-half method using odd-even items. The
Spearman-Brown formula was used for predicting the reliability of the whole test.

The first statistical step applied to the study skills test was for the purpose of determining the value of the instrument as a measuring device. There is no sense starting out with a rubber yardstick (8, p. 430). According to Bradley (7), all testing instruments should have their value as a measuring device partially determined by the reliability or consistency with which the test measures from one time to another. There are several ways to determine the coefficient of reliability. Two of the ways are the test-retest and the split-half methods.

A group of subjects is tested on two separate occasions with the same instrument in the test-retest method of finding the coefficient of reliability. Garrett (11) indicates that repetition of a test is the simplest method for determining agreement between two sets of scores. The test is given, then repeated on the same sample and the correlation computed between the two sets of scores. For the ordinary personality or ability test, according to McNemar (16), one may use either the split-half or test-retest scores to ascertain the reliability of the instrument. Since the coefficient will not be accurate if the
relative ability of the subjects changes during the interval between measurements, the split-half test method is also useful. There are several objections to the test-retest method of determining the reliability coefficient.

Some of the limitations of the test-retest method of determining the reliability coefficient are worthy of consideration. If the retest is given immediately, the ability to remember items will influence the second set of scores. Transfer effect is different from person to person. Transfer effect would be greater, according to Garrett (11), if the retest is given too soon after the first test and would tend to make the reliability coefficient too high. A delay in the time between test and retest would tend to lower the reliability coefficient. Furst (10, p. 319), states that by using the same set of items the method does not give a true indication of the adequacy and representativeness of sampling from the possible tasks within the area. Since the time element is important, Baggaley (1, p. 62) insists the time interval between test and retest should be reported along with the coefficient.

The time between test and retest was one week which is in agreement with Baggaley (1, p. 62). It is probably
better that the time interval for estimation of retest reliability should be not shorter than one week but not longer than six months. The scores of the first test were correlated with the scores of the retest to give the reliability coefficient which Cronbach (9, p. 136) calls a coefficient of stability, because it tells how stable the particular performance is. Specific-lasting characteristics were found by computation of the simple correlation. The degree of relationship between test and retest can be stated in terms of a single number (2, p. 31).

Simple correlational technique was used to determine the reliability coefficient for the test-retest method. The following formula from Garrett (11) was used:

$$r = \sqrt{\frac{N \bar{xy} - (\bar{x})(\bar{y})}{\sqrt{[N \bar{x}^2 - (\bar{x})^2][N \bar{y}^2 - (\bar{y})^2]}}$$

The computation of this statistic was verified by the Computer Center at the North Texas State University by the schema employed there to program the computer for simple correlation. The mean and standard deviation for both test and retest scores were computed at the same time. The correlation for the scores of the test-retest method was .7589. The mean for the test was 25.10 and the standard deviation was 5.97. The mean for the retest was 27.27 and the
standard deviation was 6.75. To overcome some of the limitations of the test-retest method, the split-half method was also used.

Procedures have been devised for estimating reliability from a single administration of the test. Furst (10, p. 319) says a common method is to divide the items in a single test into halves, pooling the odd-numbered items for one score and the even-numbered items for the other score. In this way the coefficient of correlation between the two halves or part scores gives an estimate of reliability of one half of the whole test. However, according to Thorndike and Hagen (21, p. 179), this value is not directly applicable to the full-length test, which is actually the instrument prepared for use. Correlating the two parts, according to Cronbach (9, p. 141), gives a coefficient of equivalence for the two half tests. To obtain the coefficient for the full test the Spearman-Brown prophecy formula is often used.

The Spearman-Brown prophecy formula for estimating reliability from two halves, taken from Garrett (11, p. 339), is as follows:
where $r_{II}$ is the reliability coefficient for the whole test. Garrett (11, p. 340) further states that the split-half method is regarded by many as the best of several methods of estimating test reliability. The mean and standard deviation were computed for each half of the test. The mean for the odd-numbered items was 13.47 and the standard deviation was 3.79. The mean for the even-numbered items was 13.92 and the standard deviation was 3.82. The simple correlation between the odd-numbered and even-numbered halves was .6606. When the Spearman-Brown prophecy formula was applied for the whole test the correlation coefficient was .7956. An additional check was used by the application of the Kuder-Richardson formula 21 from Garrett (11, p. 341). This statistical application gave a correlation coefficient of .6998 for the whole test.

Validity

The validity of the study skills test as a measure of the ability to organize, evaluate, interpret, and draw conclusions from printed material was determined by an empirical comparison of this test and several criteria measures.
In many situations for which tests are developed, some more cumbersome method of collecting information is already in use. If the existing method is considered useful in decision making, the first question in validation is whether the new test agrees with the present source of information. . . . Validation again requires an empirical comparison. Both the test and the original procedure are applied to the same subjects and the results are compared (9, p. 341).

The study skills test results were compared with the results of the sub-test scores of the **Primary Mental Abilities Test**. They are verbal meaning, number facility, reasoning, and space relations. The study skills test results were also compared with the results of the **Critical Thinking Appraisal**. At the end of the semester the results of the study skills test were compared with grade averages. Simple correlational statistics were run on all of the test results. Cronbach (9, p. 104) says there are two types of validity possible here. The type of empirical check on agreement between tests is called concurrent validity because the two sources of information are obtained at nearly the same time. The correlation coefficient could have been used to predict grade-point averages and this would have been predictive validity. However, he states, almost never does a research report tell of concurrent and predictive validity being determined at the same time. Simple
correlation was used to determine the validity coefficient for each of the test scores and the grade-point average. A correlation coefficient was computed for the grade-point averages and each of the six predictor variables which are as follows: verbal meaning, number facility, reasoning, spatial relations, critical thinking, and the study skills test. The sample used in the validity study was 120 subjects. As the size of the sample increases, the probability that the sample is like the population increases (7, p. 108).

The validity coefficient was determined. The Pearson-Product Moment Method of Correlation was used to compute the relationship between the following six predictors and study skills criterion scores:

1. PMA verbal meaning
2. PMA reasoning
3. PMA spatial relations
4. PMA number facility
5. Critical thinking
6. School grades

School grades were given the following point value for each semester hour:
Grade of A 4 points
Grade of B 3 points
Grade of C 2 points
Grade of D 1 point
Grade of F 0 point

As noted previously, the computation of statistics was verified by the Computer Center at the North Texas State University by the schema employed to program the computer for simple correlation. Since the mean and standard deviations are units to be eliminated by computation of the correlation coefficient (13, p. 206), the mean and the standard deviation were computed for each of the variables by the Computer Center.

The intercorrelation matrix for all variables included in this study is reported in Table I, page 49.

As noted in Table I, the correlation of .655 between study skills and grade-point averages is the largest, and the correlation of .262 between study skills and spatial relations is the smallest correlation where study skill is involved. However, there is a smaller correlation of .146 between grade-point average and spatial relations. The correlation between study skills and grade-point average of .655 with an N of 120 is significantly different
<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Verbal meaning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>2. Number facility</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Reasoning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Spatial relations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Study Skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Critical thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Grade-point average</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE I**

INTERCORRELATION MATRIX OF VARIABLES FOR 120 SUBJECTS
from zero at the .01 level, while the correlation of .262 between study skills and spatial relations is not significant at the .05 level. Both males and females were used as subjects in this part of the validity study. However, separate norms for males and females were determined.

The intercorrelation matrix for all variables included in this study with all male subjects in a sample of sixty-two observations is given in Table II, page 51.

As shown in Table II, for this group of sixty-two males, the most significant correlation is .603 between study skills and grade-point averages. With an N of sixty-two, the correlation of .603 is significant at the .01 level. The correlation of .245 between study skills and spatial relations is lower than it was for the total group. It is not significant at the .05 level. This is the lowest correlation recorded for this group of subjects.

The mean and standard deviation for each variable is shown in each of the separate matrices.

The intercorrelation matrix for all variables included in this study with all female subjects in a sample of fifty-eight observations is given in Table III, page 52.
TABLE II
INTERCORRELATION MATRIX OF VARIABLES FOR 62 MALES

<table>
<thead>
<tr>
<th>Variable</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Verbal meaning</td>
<td>.397</td>
<td>.415</td>
<td>.259</td>
<td>.408</td>
<td>.564</td>
<td>.247</td>
<td>20.87</td>
<td>8.77</td>
</tr>
<tr>
<td>2. Number facility</td>
<td>.652</td>
<td>.400</td>
<td>.641</td>
<td>.488</td>
<td>.485</td>
<td></td>
<td>17.90</td>
<td>5.47</td>
</tr>
<tr>
<td>3. Reasoning</td>
<td>.469</td>
<td>.552</td>
<td>.476</td>
<td>.365</td>
<td></td>
<td></td>
<td>35.70</td>
<td>10.04</td>
</tr>
<tr>
<td>4. Spatial relations</td>
<td>.245</td>
<td>.345</td>
<td>.187</td>
<td></td>
<td></td>
<td></td>
<td>33.29</td>
<td>12.90</td>
</tr>
<tr>
<td>5. Study skills</td>
<td>.522</td>
<td>.603</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26.50</td>
<td>6.18</td>
</tr>
<tr>
<td>6. Critical thinking</td>
<td>.513</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>62.62</td>
<td>8.59</td>
</tr>
<tr>
<td>7. Grade-point average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.36</td>
<td>.73</td>
</tr>
<tr>
<td>Variable</td>
<td>Variable</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------</td>
<td>------</td>
<td>--------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Verbal meaning</td>
<td>2, 3, 4, 5, 6, 7</td>
<td>20.91</td>
<td>6.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Number facility</td>
<td></td>
<td>14.29</td>
<td>5.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Reasoning</td>
<td></td>
<td>37.01</td>
<td>8.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Spatial relations</td>
<td></td>
<td>26.05</td>
<td>11.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Study skills</td>
<td></td>
<td>27.56</td>
<td>6.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Grade-point average</td>
<td></td>
<td>2.58</td>
<td>.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As noted in the preceding table, the most significant correlation for the fifty-eight female subjects was .710 between study skills and grade-point averages. This correlation was significant at the .01 level. The lowest correlation for female subjects was .213 between grade-point average and spatial relations. This was not significant at the .05 level.

Other Criteria

The study skills test which is attached as Appendix A is a series of questions designed to measure the subject's ability to read, locate information, recall information, understand instructions, organize, reason, solve problems, and use charts and graphs. This test is designed to test the subject's use rather than his knowledge of study skill. In each instance the subject is required to use his acquired skill to find the correct answer to the question. The test items have been collected from various groups of students who are interested in tests and measurements.

There are several tests currently used to predict success in school. The Watson-Glaser Critical Thinking Appraisal and the Primary Mental Abilities tests are two of them.
The Primary Mental Abilities test was chosen, because it is widely accepted as a "good" measure of general ability. The Watson-Glaser Critical Thinking Appraisal was chosen as a standardized measure of a college freshman's reasoning ability. These tests will be more fully described in the following paragraphs.

The Watson-Glaser Critical Thinking Appraisal measures five aspects of the subject's ability to think critically. They are the ability to draw sound inferences from a statement of facts, to recognize assumptions implied by a statement, to reason logically by deduction, to reason logically by interpretation, and to discriminate between strong and weak arguments. The test is in the form of printed passages that are read by the subject and then judged by the five different criteria listed above. The subject may make any of five different decisions as to the value of each question. These choices are: (1) definitely true, (2) probably true, (3) definitely false, (4) probably false, and (5) insufficient data. The norms for this test are furnished in the form of percentiles. According to the manual (22) the reliability of the test as a whole and of the separate sub-test scores have been determined by both split-half method
results in an average reliability coefficient for several groups of .84. The inter-form method for several groups gives an average reliability coefficient of .91.

The **Primary Mental Abilities** tests are designed to provide multifactored and general measures of intelligence. The two kinds of scores represent the application of the work of L. L. and Thelma G. Thurstone in the development of the theory of mental measurement. Scores on the primary mental abilities were incorporated within the framework of a single-quotient score for the entire test.

The four factors of intelligence measured by the PMA are those that appear to be most critical in schoolwork. The four primary mental abilities measured by the test are, briefly, as follows:

**V - Verbal meaning.** -- The ability to understand ideas expressed in words. In the later school years this is the most important single index of a child's potential for handling academic tasks.

**N - Number facility.** -- The ability to work with numbers, to handle simple quantitative problems, and recognize quantitative differences.

**R - Reasoning.** -- The ability to solve logical problems.
**Spatial relations.**--The ability to visualize objects and figures rotated in space and the relations between them (18).

**Subjects**

The subjects participating in this study were freshmen at the Southwest Baptist College of Bolivar, Missouri, which is a liberal arts junior college. Most of these students are graduates of Missouri high schools. They are from both rural and urban areas.

The parents of the subjects are in the lower-middle class (19, p. 447). They follow a variety of vocations. Some are skilled craftsmen, small businessmen, farmers, teachers, and ministers.

The subjects were chosen by classrooms rather than on an individual basis. Three classes of freshman English students and one class of Educational Psychology students were used in the test-retest portion of the reliability study. All three English classes were under the direction of the same teacher. For the validity study, three classes of Introductory Psychology and one class of Educational Psychology were used as subjects.
The major field of concentration of each subject was not taken into consideration because most of the students went to other colleges to finish their work toward a degree. Many of the subjects have not made a firm decision about their major field when they finish their work at this junior college.

Collection of Data

It was determined in planning this study that a sample of not less than 100 subjects should be used in both the reliability and validity studies of the test. Since there were no classes this large available, it was necessary to use several classroom groups as subjects.

The students enrolled in one section of Educational Psychology and students enrolled in two sections of English Composition were used as subjects in the reliability study. The test-retest method of determining the reliability coefficient required two administrations of the test to the same group of subjects.

The study skills test was given to subjects at a regular class period. The subjects were not informed prior to the testing period that they were to take part in the test. Before the mimeographed test questions were distributed,
the subjects were told that a study of college freshmen
was being conducted and that they had been chosen as sub-
jects. They were informed that their ability to think
and to reason was being measured. They were told that
they would be given the results of the test when it had
been scored. Cooperation of the group was good. Answer
sheets were distributed first.

The answer sheets were made by listing numbers from
one to 100 in four columns. To the right of each item
number were placed the numbers one, two, three, four.

These numbers were for the four choices a subject
could make for each item. The subjects were told to write
their names in the space provided for it. They were told
that each question on the test would have four possible
answers and they were to circle the number on their answer
sheet that corresponded to the choice they had made on the
test. Cronbach (9, p. 44) states that for efficient test-
ing subjects must follow instructions promptly and all
must do the same thing.

After the test booklets had been distributed to each
subject and placed face down on the desk, the following
instructions were given. "There are fifty questions on
this test. Each question has four answers from which you
are to choose the correct one. Choose the answer which seems to you to be the best one then circle the corresponding number for that question on your answer sheet. Make a circle around only one choice for each of the fifty questions. Make all of your marks on the answer sheet. Do not make any marks on the test booklet. Are there any questions? You have fifty minutes to complete the test. You may begin."

All of the subjects were allowed only fifty minutes to complete the test. At the end of the fifty-minute period, first the test booklets then the answer sheets were collected. The subjects were not told at this time the test would be given the next week for the retest phase of the study.

One week after the first administration of the test, it was given to the same subjects again. On the second administration a different introduction to the testing situation was given. The subjects were informed that all of the information needed for the study of college freshmen had not been secured. They were requested to do as well as possible on the test in order that the result might be compared to the result of the test they had taken the week before. They were given the same instructions
that they had received on the first administration of the test. The procedure was the same for distribution and collection of test booklets and answer sheets that it was on the first administration of the test. Most of the subjects readily accepted the new task in good spirit. The completion of this test provided the data necessary for a measure of reliability by the test-retest and split-half test methods. The reliability study had a sample of 107 subjects. A different sample of subjects was used for the study of validity.

Two tests were used as criterion measures for the development of the study skills test. The two tests were the Watson-Glaser Critical Thinking Appraisal and the Primary Mental Abilities tests. These two tests along with the study skills test were administered to a sample of 120 freshmen at the Southwest Baptist College, Bolivar, Missouri, in March, 1964. The subjects were enrolled in several different classes.

Three sections of general psychology and one section of educational psychology were used to provide the 120 subjects for the study. The tests were given in the same order to all four of the different groups.
All four sections met class on a Monday-Wednesday-Friday schedule. Monday was selected for the testing period. The Watson-Glaser Critical Thinking Appraisal was administered to each section at the regular class hour. The entire 120 subjects were given the test on the same day. The purpose was to hold constant any possible effects of motivation or boredom which might be present at different periods during the week. The test results were recorded on a tally sheet prepared for that purpose.

The tally sheet was prepared by listing the student's names down the left side on an eight and one half by eleven inch sheet of paper. Across the top of the sheet were placed the six headings for the test results. A space at the top of the sheet was left blank for the recording of the grade-point averages that could not be secured until the end of the semester. The Watson-Glaser Critical Thinking Appraisal scores were the first to be recorded.

After the Watson-Glaser Critical Thinking Appraisal was scored the results were recorded on the tally sheet in raw scores. No order was followed in recording the scores. The next test given was the Primary Mental Abilities test.
One week after the *Watson-Glaser Critical Thinking Appraisal* was given the subjects were given the *Primary Mental Abilities* test. All of the 120 subjects took the test on the same day. Monday was used for this test in order that the time effect would be the same for this test as it was for the first one. In the administration of both tests the instructions were read from the *Administrator's Manual* that is provided by the test publisher. The time limits for testing were followed and there were no unusual interruptions. Three of the 120 subjects were absent but were found and given the test on Monday afternoon. The same testing procedure was followed in testing these three subjects that was used with the larger groups. The test results were then recorded on the tally sheet under the proper headings.

Since the *Primary Mental Abilities* tests include four factors of intelligence, four different scores were recorded for each student. These four factors were verbal meaning, number facility, reasoning, and spatial relations. The next test given was the study skills test.

On Monday, one week after the administration of the *Primary Mental Abilities* test, the study skills test was
given to the same subjects. Only 113 subjects were present. It was necessary for two subjects to take the test on Tuesday. The same procedure was used with the two subjects that was used with the larger groups. All of the subjects were given the same instructions.

The subjects were told that the test items were designed to measure their ability to think and to reason. They were informed that there was only one correct answer for each of the questions, and that they would have fifty minutes in which to complete all of the items. They were then told to start. At the end of the class period they were told to stop.

The tests and answer sheets were collected. These were the same copies of the test that were used in the reliability study. The answer sheets were the same type that were used in the reliability study and were scored with the same key.

The scores from the study skills test were added to the tally sheet. This gave a total of six different test scores. The test scores on the tally sheet were the Watson-Glaser Critical Thinking Appraisal, verbal, number facility, reasoning, and spatial relations from the Primary
Mental Abilities test, and the scores from the study skills test. One more score was added to the tally sheet.

The grade-point average for each subject was furnished by the registrar's office. The grade-point average was for the hours completed by each subject at the close of the Spring Semester, 1964. Physical education grades were not included in the computation of the grade-point average. The grade-point averages were added to the tally sheet completing the collection of the data necessary for the study.

The purpose of quantifying variables is to permit a precise description of attributes and the relations among them (20). From the first examination of the pilot study to the final conclusion of the study all variables were treated according to their numerical value for descriptive purposes.
CHAPTER BIBLIOGRAPHY


The statistical treatment of the data by simple correlational methods showed that the study skills test could be used for prediction. To find the best combination of predictors, however, it was necessary to use another statistical method.

To find the best combination of predictors, a multiple correlation coefficient was computed for the total number of subjects \(N = 120\), and a multiple correlation was computed for males and females separately.

A multiple correlation was computed for the total sample of 120 subjects using the six predictors, namely verbal meaning, number facility, reasoning, spatial relations, study skills and critical thinking, and the criterion, grade-point average. Each predictor was so weighted that the correlation between the weighted sum of the predictors and the criterion was as high as possible. The multiple correlation was computed by the Computer Center at the North Texas State University using the schema designed for that purpose.
The multiple correlation involving all six predictors for the total sample of 120 subjects was .6765 which is significantly different from zero at the .01 level. According to McNemar (1), multiple correlations based on an N of less than 100 is not dependable. However, since the data are available, it is being reported. The multiple correlation for 62 males was .6650 which is significant at the .0 level. The multiple correlation for 58 females was .7381 which is significant at the .0 level.

Hypothesis one stated that the test instrument developed for this study would be reliable. This hypothesis was verified by a test-retest reliability coefficient of .7589. It was further verified by a split-half reliability coefficient of .6606. When the Spearman-Brown formula was applied, the reliability coefficient for the full test was .7956. The application of the Kuder Richardson formula 21 gives reliability coefficient of .6998.

Hypothesis two stated that the relationship between study skills as measured by the test and the six listed predictors of academic success would be positive and in the following rank order from highest to lowest:

1. School grade-point averages
2. Critical Thinking Scores
3. PMA Verbal meaning
4. PMA Reasoning
5. PMA Spatial relations
6. PMA Number facility

This hypothesis was partially verified. The relationship between study skills and the predictors was positive but did not follow the exact order of the predicted relationships and the actual rank order of the relationships as shown in Table IV below.

**TABLE IV**

ACTUAL AND PREDICTED RELATIONSHIPS OF STUDY SKILLS AND PREDICTOR VARIABLES

<table>
<thead>
<tr>
<th>Variable</th>
<th>Actual</th>
<th>Predicted</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade-point averages</td>
<td>1</td>
<td>1</td>
<td>.6551</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>2</td>
<td>2</td>
<td>.5786</td>
</tr>
<tr>
<td>Reasoning</td>
<td>3</td>
<td>4</td>
<td>.5573</td>
</tr>
<tr>
<td>Number facility</td>
<td>4</td>
<td>6</td>
<td>.5501</td>
</tr>
<tr>
<td>Verbal meaning</td>
<td>5</td>
<td>3</td>
<td>.4691</td>
</tr>
<tr>
<td>Spatial relations</td>
<td>6</td>
<td>5</td>
<td>.2625</td>
</tr>
</tbody>
</table>

The first two rank order variables were predicted but the other four were in an unpredicted order.
Hypothesis three stated that the combination of variables most significantly related to study skills was grade-point averages, critical thinking, and verbal meaning.

This hypothesis was verified for females and males when the F level of significance was computed for each separately. When the F level of significance was computed for the total sample, verbal meaning was not one of the more significant variables. In Table V, below, is shown the F level of significance for the total sample with the criterion variable, grade-point average entered at the zero level of significance.

**TABLE V**

THE F LEVEL OF SIGNIFICANCE FOR THE SIX PREDICTOR VARIABLES FOR THE SAMPLE OF 120 SUBJECTS

<table>
<thead>
<tr>
<th>Variable</th>
<th>F Level</th>
<th>Beta</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study skills</td>
<td>88.753</td>
<td>.514</td>
<td>.001</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>4.325</td>
<td>.190</td>
<td>.01</td>
</tr>
<tr>
<td>Spatial relations</td>
<td>.957</td>
<td>.106</td>
<td>.05</td>
</tr>
<tr>
<td>Reasoning</td>
<td>.561</td>
<td>.051</td>
<td>.05</td>
</tr>
<tr>
<td>Number facility</td>
<td>.213</td>
<td>.043</td>
<td>.05</td>
</tr>
<tr>
<td>Verbal meaning</td>
<td>.017</td>
<td>.130</td>
<td>.05</td>
</tr>
</tbody>
</table>
As noted in Table V, both study skills and critical thinking are greater than the .01 level of significance.

Table VI shows the F level of significance with the six predictor variables for the sample of fifty-eight female subjects.

**TABLE VI**

THE F LEVEL OF SIGNIFICANCE WITH SIX PREDICTOR VARIABLES FOR THE SAMPLE OF 58 FEMALE SUBJECTS

<table>
<thead>
<tr>
<th>Variable</th>
<th>F Level</th>
<th>Beta</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study skills</td>
<td>56.976</td>
<td>.539</td>
<td>.001</td>
</tr>
<tr>
<td>Verbal reasoning</td>
<td>3.880</td>
<td>.213</td>
<td>.01</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>.371</td>
<td>.092</td>
<td>.05</td>
</tr>
<tr>
<td>Spatial relations</td>
<td>.485</td>
<td>.081</td>
<td>.05</td>
</tr>
<tr>
<td>Number facility</td>
<td>.131</td>
<td>.381</td>
<td>.05</td>
</tr>
<tr>
<td>Reasoning</td>
<td>.016</td>
<td>.019</td>
<td>.05</td>
</tr>
</tbody>
</table>

It can be seen from Table VI that the relationship between the predictors is as predicted in hypothesis three for female subjects. Both study skills and verbal reasoning were significant beyond the .01 level. Although critical thinking was not significant, it came next in order of significance. The F level of significance for the sixty-two male subjects is shown in Table VII.
TABLE VII
THE F LEVEL OF SIGNIFICANCE FOR SIX PREDICTOR VARIABLES FOR THE SAMPLE OF 62 MALE SUBJECTS

<table>
<thead>
<tr>
<th>Variable</th>
<th>F Level</th>
<th>Beta</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study skills</td>
<td>34.348</td>
<td>.427</td>
<td>.001</td>
</tr>
<tr>
<td>Critical thinking</td>
<td>5.492</td>
<td>.334</td>
<td>.001</td>
</tr>
<tr>
<td>Verbal meaning</td>
<td>1.399</td>
<td>.146</td>
<td>.05</td>
</tr>
<tr>
<td>Number facility</td>
<td>.791</td>
<td>.155</td>
<td>.05</td>
</tr>
<tr>
<td>Reasoning</td>
<td>.234</td>
<td>.056</td>
<td>.05</td>
</tr>
<tr>
<td>Spatial relations</td>
<td>.724</td>
<td>.031</td>
<td>.05</td>
</tr>
</tbody>
</table>

It can be seen from the above table that the three variables, study skills, critical thinking, and verbal meaning are in the order predicted by hypothesis three. Only the first two have an F level of significance greater than .05.

Summary

Hypothesis one was verified by the development of a reliable instrument. The reliability coefficient was .7589 for the test-retest and .7956 for the split-half method.

Hypothesis two was partially verified since the relationship of the six predictor variables was positive and
the first two were in the order predicted. The validity coefficient was larger for grade-point averages and study skills than it was for study skills and any other variable.

Hypothesis three was partially verified. By the use of McNemar's (1, p. 175) formula for multiple correlation coefficient for three variables,

\[ r_{123} = \sqrt{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}} \]

it can be shown that the two variables which were predicted to be most related to study skills yield a larger coefficient than any two variables taken alone. The simple correlations were as follows: study skills and grade-point averages .65, study skills and critical thinking .57, grade-point averages and critical thinking .49. The multiple correlation for the three variables was .70. Grade-point averages and critical thinking were the two variables predicted to be the most significantly related. Although verbal meaning would have increased the multiple correlation coefficient, it was not included because the F level was less than .05.
Item Analysis

The test was subjected to item analysis and the following statistics were computed for each item: per cent correct response, per cent discrimination, and per cent efficiency. The per cent correct response was determined by computing the per cent of students who answered the item correctly. The per cent of discrimination was determined by subtracting the bottom 27 per cent of the students from the top 27 per cent in terms of per cent correct response. The per cent efficiency was determined by means of the Flanagan table. For the result of the item analysis see Appendix B.
CHAPTER BIBLIOGRAPHY

The purpose of this study was the development of an instrument to determine the study skills of college freshmen. In the development of the testing instrument it was necessary to determine its reliability and validity. The reliability was determined by both the split-half and test-retest method for determining the reliability coefficient. The validity coefficient was obtained by simple correlation to determine the extent of the relationship between this instrument and six other predictor variables of college success. The predictive validity was obtained by multiple correlation with five other variables to determine which of the six variables had the most predictive value. The criterion variable was school grade-point averages.

The following hypotheses were investigated by statistical analysis of the data that were collected.

1. The test instrument developed for this study will be a reliable study skills test.
2. The relationship between study skill as measured by the test to be developed and the six listed predictors of academic success will be positive and in the following rank order of magnitude from highest to lowest:

- a. Grade-point averages
- b. Critical thinking scores
- c. PMA Verbal meaning
- d. PMA Reasoning
- e. PMA Spatial relations
- f. PMA Number facility

3. The combination of variables most significantly related to study skill will be grade-point averages, critical thinking, and verbal meaning.

The subjects used in this study were from two different schools. The subjects used in the pilot study were students at the North Texas State University, Denton, Texas. The subjects used for the reliability and validity study were students at the Southwest Baptist College, Bolivar, Missouri.

The subjects used in the pilot study were all freshmen enrolled in an orientation course. There were sixty students enrolled in three sections of Education 161. In the reliability study one section could not be given the retest
because of a scheduling problem. Only thirty of the students completed both the test and retest. On the validity portion of the pilot study there were only thirty subjects who completed both the study skills test and the Watson-Glaser Critical Thinking Appraisal.

The subjects used for the reliability and validity study were freshmen students at the Southwest Baptist College enrolled in three sections of general psychology, one section of educational psychology, and three sections of English composition. They were second semester freshmen enrolled in the Spring, 1964, semester. The subjects were divided into two groups by classes. A sample of 107 subjects were used for the reliability study. No section was divided. Each subject in the same section was used for the same part of the study. A sample of 120 subjects were used for the validity study. There were fifty-eight females and sixty-two males.

The pilot study was made in December, 1963, and January, 1964. The study involved only about two weeks. The reliability and validity study was made in February and March, 1964. The grade-point averages were for the Spring semester, 1964.
The tests that were used in the study were the \textbf{Primary Mental Abilities} tests (revised 1962) which were designed to provide a multifactored measure of intelligence and the \textbf{Watson-Glaser Critical Thinking Appraisal} which was designed to provide problems and situations which require the application of some of the important abilities involved in critical thinking. The scores of the separate sub-tests of the \textbf{Primary Mental Abilities} tests were used. They were verbal meaning, number facility, reasoning, and spatial relations. Only the total score of the \textbf{Watson-Glaser Critical Thinking Appraisal} was used. This total score included the results of five sub-tests. The five sub-tests were the following: inference, recognition of assumptions, deduction, interpretation, and evaluation of arguments. The split-half reliability coefficient for the \textbf{Watson-Glaser Critical Thinking Appraisal} averages about .90 while the inter-form reliability coefficient is about .81 (2, p. 9). The norms for the \textbf{Primary Mental Abilities} tests were based on a nationwide standardization program conducted in the spring of 1962. A total of seventy-three schools in thirty-nine school systems with an enrollment of 32,708 students comprised the sample on which the test was standardized (1, p. 25).
Conclusions

In view of the results of the investigation and within the limits of the study, the following conclusions appear to be justified.

1. The test developed by this study is a reliable study skills test with a reliability coefficient of .75. Although some of the study skills tests cited in this study have a reliability coefficient larger than this, a coefficient of .75 warrants confidence in this test.

2. The test developed by this study has a validity coefficient larger than the coefficient of any of the predictor variables when compared to grade-point averages, the criterion measure. The test is a more valid measure by which grade-point averages can be predicted than any of the following predictors: critical thinking, verbal meaning, reasoning, number facilities or spatial relations.

3. A combination of predictors is significantly better than any single predictor of grade-point averages. Study skills, critical thinking and verbal meaning are the best combination of predictors when related to the following variables: study skills, critical thinking, verbal meaning, reasoning, spatial relations and number facilities.
4. The combination of predictor variables used in this study accounted for 45 per cent of the variance in predicting grade-point averages. It can be concluded that no single predictor or combination of predictors has yet been found that will account for all of the variance in predicting academic success.

Recommendations

The evidence presented in this study suggested other areas of investigation. The following recommendations are made for such investigation.

1. A combination of tests needs to be developed to account for that portion of variance unaccounted for by this combination of predictors. The combination should include study skills, critical thinking and verbal meaning which were the best predictors of grade-point averages in this study.

2. A much larger sample of subjects should be used in any future study. The 120 subjects used in this validity study should be considered as an absolute minimum for any study of validity. The sample of subjects should be taken from several different colleges to give the results wider application.
3. Four of the items on the test showed no predictive value and should be removed from the test. They were items number 9, 36, 43, and 50. These items indicate a further refinement by item analysis was needed to make the test a more effective predictor of academic success.

4. Some incentive should be used to insure the cooperation of the subjects. Since students are not anxious to be subjects in tests that involve the effort required by study skill tests, some motivation should be used. Parallel forms should be used to determine reliability as this would prevent the resentment the student has when he is asked to take the same test a second time in the test-retest method of determining the reliability coefficient.
CHAPTER BIBLIOGRAPHY


APPENDIX A

STUDY SKILL TEST

1. Antonyms are words where meanings are opposite. Select the pair of words that is not opposite in meaning.

1. occasional-frequent
2. permit-forbid
3. expedition-journey
4. happy-sad

2. In order of arrangement in a book the last printed material is

1. index of references
2. index of subjects
3. index of tables
4. index of unfamiliar words

3. The eu in eulogy means

1. short funeral oration
2. good
3. a speech for a friend
4. one's last statement

4. The prefixes in unilateral and monogram both denote

1. agreement
2. direction
3. writing
4. amount

5. The abbreviation circ means

1. which see
2. about
3. in the same place
4. compare
6. The root **jectus** means to throw. Which of the following words means to cast down or discourage?

1. projected
2. rejected
3. subjected
4. dejected

Read the following paragraph carefully as the next four questions are related to it.

1. There is some reason to believe that some behavior patterns were established in social relations with family and friends in the first few years of life. 2. Emotional responses, the way one reacts to other people and situations, are learned responses. 3. They are learned through experiencing contacts and situations in life. 4. A person may have certain emotional attitudes that were learned between the ages of two and five. 5. He may still be responding with immature patterns he established at an early age. 6. When one carries the behavior of childhood into adult life he is emotionally immature.

7. The key idea of the paragraph is

1. maturity is essential to growth
2. social relations relate to people
3. emotions are learned
4. the importance of early life

8. One could form an outline of sentences

1. 1-5-6
2. 2-4-6
3. 4-5-6
4. 1-3-6

9. A broad definition is found in sentence

1. 2
2. 1
3. 5
4. 6
10. The paragraph states

1. there is some emotional hangover in adult life
2. age two to five determines adult emotional stability
3. there could be some emotional problems stemming from childhood
4. childhood emotional responses determine adult behavior

11. To move from one idea to another writers use

1. connectives
2. transition
3. conjunctions
4. conclusions

12. The sum of the numbers from one through nine is equal to

1. the product of the middle number and the last number
2. the sum of the middle number and the next to last number
3. the square of the last number divided by two
4. the answer is not given

13. Two times what number, is equal to the square root of 256?

1. 32
2. 4
3. 16
4. 8

14. Concentration is to recall as:

1. the telephone is to talking
2. a lake is to boating
3. air is to breathing
4. a rod and reel is to fishing

15. Tension is the degree of being strained to stiffness. A speedometer is an instrument for indicating speed. Tensiometer means
1. the rate of speed required to build up tension
2. a unit of measure of tension
3. a device to cause tension
4. a device to determine tautness

16. Hemisphere is a half sphere. Photoelectric cells react to light. Photosphere means

1. an electric camera
2. half of a light globe
3. the luminous envelope of the sun
4. a sphere produced by electricity

17. Which of the following statements contain one or more emotionally-toned words?

1. The late 1940's and the decade of the 1950's witnessed economic progress.
2. The last half century was a time of rugged individualism. Gigantic individual fortunes were made and lost overnight.
3. This was also a period of gross economic injustices.
4. 1-2-3 above all contain emotionally-toned words.

Below is a picture of a card that you might find in the card catalogue of a library. Following the card there are two questions that can be answered from information given on the card.
18. The number in the upper left corner of the card is

1. the call number
2. the Dewey Decimal classification number
3. the Library of Congress classification
4. both 1 and 2


1. true
2. possibly true
3. false
4. possibly false

In the next four items match the underlined words in the phrases with words that are opposite in meaning.

20. a slovenly person
1. happy

21. performing with great frivolity
2. awkwardness

22. a woebegone expression
3. neat

23. playing with dexterity
4. seriousness

In 1950 more radios were made than at any other time in history. A total of 1,750,000 sets were made.

Using this information applied to the graph, answer the following questions. The next three questions are related to the graph.

24. The total number of radios produced in 1920, 1930, and 1940 were

1. 1/2 of the 1950 figure
2. 1-1/12 more than in 1950
3. 7/8 of the 1950 figure
4. about the same as the 1950 figure

25. How would you account for this production figure in 1930?

1. the dollar had lost its purchasing power
2. the invention of television
3. radios had not become popular by that time
4. the poor distribution of purchasing power

26. What per cent of total production was reached in 1936?

1. 25%
2. less than 25%
3. about 35%
4. not enough information to tell

27. The commander-in-chief of the United States Army in World War II was

1. Eisenhower
2. Marshall
3. Patton
4. none of these men

28. If the hour hand is pointing toward three and the minute hand toward four, the angle between them is

1. 30 degrees
2. 12 degrees
3. 20 degrees
4. 60 degrees

29. If a U.S. Representative represents only the people of his district, then each of the two U.S. Senators must represent

1. 1/2 the people of his state
2. all the people of his state
3. the people of three congressional districts
4. all the people of the United States

30. In what age range does the greatest per cent increase in vocabulary take place?
1. 6-8
2. 14-18
3. 19-25
4. 1-3

31. "It is not well for a man to pray cream and to live skim milk."

1. durability
2. dairying
3. conscientiousness
4. consistency

32. "Success is getting what you want. Happiness is wanting what you get."
Which is the best inference?

1. success also depends on wanting what you get.
2. happiness also depends on getting what you want.
3. success and happiness occur together.
4. success and happiness may occur separately.

33. In order to determine the truth of a television commercial one should

1. seek other areas of information
2. take the word of the announcer
3. try the product for one's self
4. ask someone who has tried the product

34. Prediction is based on

1. ideas concerning basic facts
2. ideas and observations
3. ideas, methods, and observations
4. ideas, methods, observations, and experiments

35. To select important information from a mass of available data, one should

1. get an assistant to aid you in judging
2. use some criterion to evaluate information
3. learn the entire field and select the best
4. consider only facts as important
36. When the first question on a ten-question examination is most difficult and must be considered carefully you should answer it before you read question number two.

1. true
2. false
3. probably true
4. probably false

37. John makes low grades because he seldom studies until just before a test, at which time he crams late at night, losing sleep and energy. Which statement would best apply to his situation?

1. speak softly, carry a big stick
2. look before you leap
3. haste makes waste
4. procrastination is a thief of time.

The next six questions are based on the table below

**BUSINESS BAROMETERS**

(1930-52 = 100)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wholesale trade</td>
<td>134.0</td>
<td>133.5</td>
<td>125.8</td>
<td>129.2</td>
<td>125.7</td>
<td>+2.9</td>
</tr>
<tr>
<td>Total Employment</td>
<td>138.1</td>
<td>139.5</td>
<td>136.5</td>
<td>138.5</td>
<td>134.8</td>
<td>+2.7</td>
</tr>
<tr>
<td>Unemployment</td>
<td>192.6</td>
<td>176.7</td>
<td>198.6</td>
<td>192.7</td>
<td>172.1</td>
<td>+11.8</td>
</tr>
<tr>
<td>Retail Trade</td>
<td>151.2</td>
<td>156.6</td>
<td>134.1</td>
<td>143.4</td>
<td>135.8</td>
<td>+5.4</td>
</tr>
<tr>
<td>Building Permits</td>
<td>152.8**</td>
<td>148.3</td>
<td>152.3</td>
<td>159.0</td>
<td>181.0</td>
<td>-12.2</td>
</tr>
<tr>
<td>R.R. Freight Tonnage</td>
<td>102.5</td>
<td>107.4</td>
<td>97.9</td>
<td>103.0</td>
<td>99.8</td>
<td>+3.1</td>
</tr>
<tr>
<td>Life Insurance Activity</td>
<td>338.7</td>
<td>283.2</td>
<td>260.2</td>
<td>258.6</td>
<td>199.7</td>
<td>+29.2</td>
</tr>
<tr>
<td>Petroleum Production</td>
<td>86.8**</td>
<td>87.7</td>
<td>87.1</td>
<td>88.3</td>
<td>83.4</td>
<td>-0.2</td>
</tr>
<tr>
<td>Gas Consumption</td>
<td>106.3</td>
<td>111.8</td>
<td>107.2</td>
<td>110.2</td>
<td>116.2</td>
<td>-5.3</td>
</tr>
</tbody>
</table>

*Represents the per cent of change of the cumulative index of January-September, 1959, from January-September, 1958.

**Estimated.
38. The figures for retail trade are expressed in:

1. thousands of dollars
2. tens of thousands of dollars
3. millions of dollars
4. none of these figures

39. The decline in petroleum production was greater between the first nine months of 1958 and the comparable period of 1959 than between 1950-52 and the first nine months of 1959.

1. true
2. false
3. probably true
4. probably false

40. In how many of the nine types of business activity listed in the table was the business situation more favorable in the first nine months of 1959 than in the years 1950-52?

1. 5
2. 6
3. 8
4. 9

41. The statement (1950-52) above the table indicates:

1. Business activity in 1950-52 forms the base from which the figures given in the table are derived.
2. Business activity in 1950-52 forms the base from which the figures given in the table, except those in the last column, are derived.
3. Business activity in 1950-52 forms the base from which the figures given in the table, except those in the last line, are derived.
4. Business activity in 1950-52 was 100 per cent greater than in 1959.

42. The percentages given in the right side of the table are based upon business activity in all the months of 1958 and 1959.
1. true
2. false
3. probably true
4. can't tell

43. How much confidence can the reader have in the accuracy of the figures given in this table?

1. complete confidence
2. a great deal
3. none
4. can't tell

44. A column in a table is

1. a horizontal array of data
2. a vertical array of data
3. the same as a line of data
4. a diagonal array of data

45. A camel has been defined as a horse put together by a committee. Which statement has the same meaning?

1. if at first you do not succeed try, try again
2. making a mountain out of a mole hill
3. the proof of the pie is in the eating
4. too many cooks spoil the broth

46. Sally was a girl of average achievements, yet she constantly chided others for being average in their achievement. Which statement has the same meaning?

1. the pot calling the kettle black
2. all's well that ends well
3. pretty is as pretty does
4. do unto others as you would have them do unto you

47. The entire universe is in a state of flux. Change is the order of the day. This idea is best expressed by

1. still water runs deep
2. you can't step in the same river twice
3. evolution is a figment of the imagination
4. oil and water do not mix
48. "Still water runs deep" implies

1. blood is thicker than water
2. potential is not measured by noise
3. skimming through the material
4. read only the amount of material you can understand

49. In reference to reading, "Separating the cream from the milk,"

1. reading the best books
2. dividing reading material into sections
3. skimming through the material
4. read only the amount of material you can understand

50. "The more you strain to appear natural, the more you appear odd," implies that

1. people must strain to be natural
2. to appear natural be yourself
3. it is difficult to appear unnatural
4. it is easy to appear unnatural.
## APPENDIX B

### ITEM ANALYSIS

<table>
<thead>
<tr>
<th>Item No.</th>
<th>% Right in Top 27%</th>
<th>% Right in Bottom 27%</th>
<th>Difficulty Index</th>
<th>Validity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.84</td>
<td>.60</td>
<td>.72</td>
<td>.36</td>
</tr>
<tr>
<td>2</td>
<td>.36</td>
<td>.21</td>
<td>.28</td>
<td>.23</td>
</tr>
<tr>
<td>3</td>
<td>.55</td>
<td>.27</td>
<td>.41</td>
<td>.33</td>
</tr>
<tr>
<td>4</td>
<td>.76</td>
<td>.33</td>
<td>.54</td>
<td>.43</td>
</tr>
<tr>
<td>5</td>
<td>.66</td>
<td>.36</td>
<td>.51</td>
<td>.32</td>
</tr>
<tr>
<td>6</td>
<td>.39</td>
<td>.39</td>
<td>.39</td>
<td>.00</td>
</tr>
<tr>
<td>7</td>
<td>.45</td>
<td>.21</td>
<td>.33</td>
<td>.26</td>
</tr>
<tr>
<td>8</td>
<td>.57</td>
<td>.45</td>
<td>.51</td>
<td>.16</td>
</tr>
<tr>
<td>9</td>
<td>.39</td>
<td>.54</td>
<td>.46</td>
<td>-.16</td>
</tr>
<tr>
<td>10</td>
<td>.70</td>
<td>.51</td>
<td>.60</td>
<td>.22</td>
</tr>
<tr>
<td>11</td>
<td>.84</td>
<td>.30</td>
<td>.37</td>
<td>.55</td>
</tr>
<tr>
<td>12</td>
<td>.70</td>
<td>.33</td>
<td>.51</td>
<td>.36</td>
</tr>
<tr>
<td>13</td>
<td>.72</td>
<td>.24</td>
<td>.48</td>
<td>.50</td>
</tr>
<tr>
<td>14</td>
<td>.52</td>
<td>.33</td>
<td>.42</td>
<td>.23</td>
</tr>
<tr>
<td>15</td>
<td>.63</td>
<td>.21</td>
<td>.42</td>
<td>.42</td>
</tr>
<tr>
<td>16</td>
<td>.52</td>
<td>.12</td>
<td>.32</td>
<td>.48</td>
</tr>
<tr>
<td>17</td>
<td>.33</td>
<td>.24</td>
<td>.28</td>
<td>.15</td>
</tr>
</tbody>
</table>
ITEM ANALYSIS—Continued

<table>
<thead>
<tr>
<th>Item No.</th>
<th>% Right in Top 27%</th>
<th>% Right in Bottom 27%</th>
<th>Difficulty Index</th>
<th>Validity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>.70</td>
<td>.30</td>
<td>.50</td>
<td>.40</td>
</tr>
<tr>
<td>19</td>
<td>.84</td>
<td>.72</td>
<td>.78</td>
<td>.20</td>
</tr>
<tr>
<td>20</td>
<td>.94</td>
<td>.30</td>
<td>.62</td>
<td>.63</td>
</tr>
<tr>
<td>21</td>
<td>.81</td>
<td>.18</td>
<td>.49</td>
<td>.62</td>
</tr>
<tr>
<td>22</td>
<td>.84</td>
<td>.55</td>
<td>.69</td>
<td>.36</td>
</tr>
<tr>
<td>23</td>
<td>.84</td>
<td>.27</td>
<td>.55</td>
<td>.58</td>
</tr>
<tr>
<td>24</td>
<td>.63</td>
<td>.36</td>
<td>.49</td>
<td>.30</td>
</tr>
<tr>
<td>25</td>
<td>.60</td>
<td>.42</td>
<td>.51</td>
<td>.23</td>
</tr>
<tr>
<td>26</td>
<td>.76</td>
<td>.52</td>
<td>.64</td>
<td>.30</td>
</tr>
<tr>
<td>27</td>
<td>.57</td>
<td>.39</td>
<td>.48</td>
<td>.20</td>
</tr>
<tr>
<td>28</td>
<td>.60</td>
<td>.12</td>
<td>.36</td>
<td>.53</td>
</tr>
<tr>
<td>29</td>
<td>.63</td>
<td>.48</td>
<td>.55</td>
<td>.16</td>
</tr>
<tr>
<td>30</td>
<td>.63</td>
<td>.21</td>
<td>.42</td>
<td>.42</td>
</tr>
<tr>
<td>31</td>
<td>.66</td>
<td>.33</td>
<td>.49</td>
<td>.35</td>
</tr>
<tr>
<td>32</td>
<td>.66</td>
<td>.33</td>
<td>.49</td>
<td>.35</td>
</tr>
<tr>
<td>33</td>
<td>.57</td>
<td>.42</td>
<td>.49</td>
<td>.20</td>
</tr>
<tr>
<td>34</td>
<td>.55</td>
<td>.39</td>
<td>.47</td>
<td>.16</td>
</tr>
<tr>
<td>35</td>
<td>.73</td>
<td>.43</td>
<td>.60</td>
<td>.30</td>
</tr>
<tr>
<td>36</td>
<td>.52</td>
<td>.57</td>
<td>.54</td>
<td>-.05</td>
</tr>
<tr>
<td>37</td>
<td>.72</td>
<td>.48</td>
<td>.60</td>
<td>.26</td>
</tr>
<tr>
<td>Item No.</td>
<td>% Right in Top 27%</td>
<td>% Right in Bottom 27%</td>
<td>Difficulty Index</td>
<td>Validity Index</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------</td>
<td>-----------------------</td>
<td>------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>38</td>
<td>.72</td>
<td>.55</td>
<td>.63</td>
<td>.20</td>
</tr>
<tr>
<td>39</td>
<td>.52</td>
<td>.27</td>
<td>.39</td>
<td>.30</td>
</tr>
<tr>
<td>40</td>
<td>.70</td>
<td>.46</td>
<td>.59</td>
<td>.23</td>
</tr>
<tr>
<td>41</td>
<td>.76</td>
<td>.48</td>
<td>.62</td>
<td>.32</td>
</tr>
<tr>
<td>42</td>
<td>.63</td>
<td>.18</td>
<td>.40</td>
<td>.48</td>
</tr>
<tr>
<td>43</td>
<td>.45</td>
<td>.52</td>
<td>.48</td>
<td>-.10</td>
</tr>
<tr>
<td>44</td>
<td>.91</td>
<td>.45</td>
<td>.68</td>
<td>.54</td>
</tr>
<tr>
<td>45</td>
<td>.84</td>
<td>.39</td>
<td>.61</td>
<td>.50</td>
</tr>
<tr>
<td>46</td>
<td>.84</td>
<td>.66</td>
<td>.75</td>
<td>.26</td>
</tr>
<tr>
<td>47</td>
<td>.72</td>
<td>.39</td>
<td>.55</td>
<td>.36</td>
</tr>
<tr>
<td>48</td>
<td>.91</td>
<td>.60</td>
<td>.75</td>
<td>.38</td>
</tr>
<tr>
<td>49</td>
<td>.81</td>
<td>.36</td>
<td>.58</td>
<td>.45</td>
</tr>
<tr>
<td>50</td>
<td>.88</td>
<td>.96</td>
<td>.92</td>
<td>-.11</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY

Books


---


---


---


---


---


---


---


---


---


---


---


---


Wilcox, G. W., Basic Study Skills, Boston, Allyn and Bacon, Inc., 1958.


Wrenn, C. G., Practical Study Aids, Stanford, California, Stanford University Press, 1957.

Articles

Barbe, W. B., "The Effectiveness of Work in Remedial Reading at the College Level," Journal of Educational Psychology, XLIII (July, 1952), 229-237.


Kilby, R. W., "The Relation of a Remedial Reading Program to Scholastic Success in College," Journal of Educational Psychology, XXXVI (January-December, 1945), 513-534.


Kraner, R. E., "A Comparison of Two Methods of Listening Instruction," Reading Newsletter, XXXIII (September, 1964), 1-12.

Laroe, Willetta R., "How to Study," Grade Teacher, LXXI (October, 1953), 56.


McQueen, R., "Diagnostic Reading Scores and College Achievement," Psychological Reports, III (May, 1957), 627-629.


Manuals of Directions

Contemporary Tests of Scholastic Progress, Monterey, California Test Bureau, 1962.


Otis Quick-Scoring Mental Ability Gamma Test, New York, World Book Company, 1939.


Study-Habits Inventory, Stanford, California, Stanford University, 1941.

