

T
312

PROPOSALS FOR
A PREDICTIVE TEST OF
ACADEMIC POTENTIAL FOR GRADUATE WORK
IN EDUCATION AT A.U.B.

By
RASHID AHMED SHAH

A Thesis

Submitted in Partial Fullfillment of the Requirements
For the Degree of Master of Arts
In the Education Department of the
American University of Beirut
Beirut, Lebanon

June, 1960

GRADUATE APTITUDE TEST

SHAH

ACKNOWLEDGEMENTS

I wish to acknowledge my heartfelt thanks and deep gratitude to Dr. F.R. Korf, Acting Chairman of the Department of Education, American University of Beirut, under whose advisorship and guidance the present research has been carried out. His encouragement, suggestions, and kind help throughout the present research, particularly in the initial planning of this study, in sending requests to graduate students who were to take the test, in criticizing test items, and in the statistical analysis of the tryout data, enabled me to fulfill this undertaking within the limited time available.

Professor J.M. Katul and Dr. Naim Atiyeh, members of my Thesis Committee, deserve my gratitude and appreciation for their constructive criticisms during the defense of the outline, and also for their ready help during the study.

I am also indebted to Dr. L.P. Cajoleas and Dr. R.W. Burnett. The former for his suggestions and help, the latter for his encouragement.

I desire to thank all the professors and graduate students in the Education Department, American University of Beirut, who helped me in the analysis of the term Academic Potential.

I am grateful to my fellow students for their cooperative attitude and for offering their valuable time to take the test.

Special thanks are due to some of my Pakistani friends who helped in scoring the test forms. I also desire to thank my Arab friends for their encouragement.

ABSTRACT

I N T R O D U C T I O N

Due to the increasing ratio of applicants to available spaces, predictive instruments are being devised to serve as entrance or qualifying examinations for many schools and colleges. Not much work however, has been done on the prediction of academic success at the graduate level.

The present study has attempted to construct a test on the basis of an analysis of the A.U.B. graduate program in education; it is hoped that it will predict academic potential for graduate work in education at the American University of Beirut.

Need for this Study

Records of graduate students in the Education Department show that graduate students often find substantial academic difficulties and as a result, some of them withdraw and some fail. These failures result in lowering the standard of work, and in maladjustments on the parts of failing students.

METHOD OF STUDY

Definition of the term Academic Potential

For the purpose of the present study, professors and graduate students in the Department of Education were interviewed as to the

abilities, skills and knowledges required for success in graduate work in Education.

Objectives of the Test in Behavioral Terms

Each of the significant abilities, knowledges and skills isolated from the interviewees' reports and the researcher's background of graduate work in Education was translated into specific behaviors to be measured, and the test constructed accordingly. It was administered to fifty subjects.

ANALYSIS OF THE TRYOUT DATA

Reliability of the Test

The reliability coefficient of the tryout test is $+0.851$.

Selection and Revision of Test Items

On the basis of statistical analysis of tryout data, twenty-six items were discarded. Fifty items were retained for the revised test. Some of these were edited by deleting their one or two non-functioning or non-discriminating options or by replacing them with presumably better ones. Those items that were discarded included negatively-discriminating, non-discriminating and non-functioning answers or misleads.

Conclusions and Recommendations

The statistical analysis showed that, on the whole, the tryout test tended to be on the easy-side for the examinees. The tryout has

been found to give a reliable measure, as inferred from the reliability coefficients of $+.85$. The removal of some items and the editing of others will probably raise appreciably the reliability of the revised test and also make the average item difficulty of the revised test more in accord with good testing practice.

Recommendations are offered for undertaking a validity study before the test is actually used for prediction purposes.

been found to give a reliable measure, as inferred from the reliability coefficients of $+0.85$. The removal of some items and the editing of others will probably raise appreciably the reliability of the revised test and also make the average item difficulty of the revised test more in accord with good testing practice.

Recommendations are offered for undertaking a validity study before the test is actually used for prediction purposes.

TABLE OF CONTENTS

Chapters	Page
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
LIST OF TABLES	x
I. INTRODUCTION	1
Need for this study... ..	2
Achievement Versus Aptitude Tests... ..	5
Limitations of the Study	6
II. METHOD OF STUDY	
Introduction	9
Definition of the Term Academic Potential	10
Objectives of the Test in Behavioral Terms	12
Construction and Selection of Items ...	14
The nature of the sample and the testing procedure	18
Scoring	21
III. ANALYSIS OF THE TRYOUT DATA	
The Purpose of the Analysis	24
The Analysis - Procedure	24
Odd vs. even items	28
Item Difficulty	31
Discriminative Power of the Item	31
Reliability of the Test	33

Selection and Revision of Test Items	36
The Revised Test	39
Conclusions and Recommendations	40
APPENDIX A: The Tryout Test with Item-Analysis Results ...	43
APPENDIX B: The Revised Form	60
APPENDIX C: <u>Steps Leading to the Degree of Master of Arts for</u> <u>Students from other Colleges and Universities</u> <u>Entering the Department of Education, American</u> <u>University of Beirut</u> (Department of Education, American University of Beirut, July, 1959) ...	76
WORKS CITED	79
SOURCES CONSULTED	80

LIST OF TABLES

Table	Page
I. Data regarding graduate students in the Department of Education during 1956 - 1959	3
II. Distribution of weight over objectives of the test...	16
III. Frequency distribution of time required for the tryout test	20
IV. Frequency distribution of scores on the tryout test...	23
V. Item-analysis data for items, 1, 2, 10, 16, 33, and 62 in the tryout test	26
VI. Distribution of scores on subtests X and Y of the tryout test	30
VII. Measures of level of difficulty (D) and discriminating significance (Chi) of individual items in the tryout test	32

CHAPTER ONE

INTRODUCTION

Psychological tests are now extensively used in military and industrial work to select appropriate personnel in order to create efficiency and smoothness in work. The limited facilities of educational institutions and training programs, and the vast number of applicants have made it necessary to use aptitude tests to select candidates who will most probably succeed in training work.

In recent years, scholastic aptitude tests have become increasingly popular in high schools and colleges. They are specialized tests that are developed on the basis of analysis of the school programs. These tests attempt to predict success in high school and college work. The educator is coming to realize that some sort of instrument that will foretell success of candidates should be used as an entrance or qualifying examination for a training program so that the proportion of failures may be held to a minimum.

Very few research studies on the prediction of academic success have been conducted on the graduate level. According to Travers, "Of all the major areas in which work has been done on the prediction of academic success, least work of quality has been undertaken in the area of predicting success in graduate school."¹ If standardized measures are

¹Robert M.W. Travers, Educational Measurement (New York: Macmillan, 1955), p. 394.

used they usually do not work as satisfactorily as an instrument developed out of the needs of the institution in which it is to be used. Often, it is difficult to find a ready-made test that measures what the job or the training program requires.

The present study aims at making a test on the basis of an analysis of the graduate program in Education at the American University of Beirut. It is hoped that it will predict academic success in graduate work in Education at the American University of Beirut.

Need for this Study

Every year students from various outside countries apply for admission to Graduate work leading to M.A. degree in the Department of Education, American University of Beirut. These applicants are accepted as full-fledged graduate students after they spend one semester and then pass a qualifying examination - an examination, consisting of written and oral parts, conducted by the Department of Education to determine the student's ability "to undertake independent research and complete successfully the graduate work leading to the Master of Arts degree,"² and also to determine "the number of credit hours and the minimum length of residence required for the award of the M.A. degree."²

In addition to such students from outside institutions, there

²Steps Leading to the Degree of Master of Arts for Students from other Colleges and Universities Entering the Department of Education, American University of Beirut. (Department of Education, American University of Beirut, July, 1959)

For further details, see copy in Appendix C.

are also undergraduates of the University itself, who are accepted as graduate students in education after they take their B.A. degrees.

It has been found that graduate students often find substantial academic difficulties during the course of their work. The data in Table I is obtained from the records of 155 past and present graduate students in the Department of Education.

TABLE I
DATA REGARDING GRADUATE STUDENTS IN EDUCATION
DURING 1956 - 1959.

Total Number of graduate students enrolled.	Number of graduate students who obtained M.A. degree.	Number of graduate students who found substantial academic difficulties		Number of students who are continuing work for the M.A. degree, or who withdrew due to sickness, their government's requirements, ^{or} unknown reasons.
		Withdrew	Failed	
155	56	4	13	82
		17		

Table I shows that out of 155 students who registered for graduate work in education during the academic period 1956-1959, fifty-six (36.1 percent) obtained the M.A. degree whereas seventeen (11 percent) found substantial academic difficulties in work leading to the M.A. degree in Education. Thus the ratio of those who obtained the M.A. degree to those who found substantial academic difficulties was approximately 7:2, i.e., for every seven successful graduate students there were two who did not complete work leading to the M.A. degree

due to academic difficulties. Out of these seventeen students, thirteen failed to earn the M.A. degree at all, whereas four withdrew during the first and second semesters of graduate work. The data in Table I also show that the remaining eighty-two students include those who are pursuing their studies for the M.A. degree, or who withdrew due to sickness, their government's requirements or unknown reasons.

Reporting about the first group of Pakistani I.C.A.-sponsored students at A.U.B., Nasim Ahmad writes:

In the field of education, out of eighteen accepted candidates, only nine were able to continue through to the M.A. degree. Out of the eight candidates who could not continue through all of the M.A. work three held [previous Pakistani] B.A. or M.A. degrees (without the B.T. degree) and failed to qualify even for the Normal Diploma.⁵

These failing students waste their time and the time of the Department. Thereby, more gifted ones lose the opportunity to get training. Besides, these failures lower the average quality of students' work and generally give rise to frustrations and maladjustments among the failing students. To avoid this situation, the development of a predictive instrument is desirable, so that, on testing candidates, a higher proportion of potentially successful students may be selected for graduate work.

From the data in Table I, it can be inferred that the present techniques of the Department to select or recommend students for graduate work do not produce highly dependable estimates of the abilities of

⁵ Nasim Ahmad, Problems of Pakistani I.C.A. Students at A.U.B., M.A. Thesis, American University of Beirut, 1959, p. 1.

candidates from outside universities to do graduate work. Hence, there is the need to attempt to develop a more adequate instrument to replace or supplement the present qualifying examinations.

It is felt that the same test may be able to fill another need, namely helping to decide on the admission of A.U.B. undergraduates to the M.A. program.

Achievement versus Aptitude Tests

The difference between an aptitude test and an achievement test is that the former measures what the individual could learn, whereas the latter gives an estimate of what the individual has learned.

As the present test attempts to measure the student's ability in independent research, in logical thinking, in comprehension, in critical analysis etc., it is felt that it may be useful in selecting promising candidates for graduate work. In this context, the present test is an aptitude test. According to Hull, ". . . an aptitude test is a test designed to discover what potentiality a given person has for learning some particular vocation or acquiring some particular skill."⁴ In short, the present test is designed to measure student's aptitude for graduate work. According to F.S. Freeman, ". . . when we speak of an individual's aptitude for a given type of activity, we mean the capacity to acquire proficiency under appropriate conditions; that is, his potentialities at present, as revealed by his performance on selected tests which have

⁴C.L. Hull, Aptitude Testing, (New York: World Book Co., 1928), as quoted by C.E. Skinner, Essentials of Educational Psychology (Englewood Cliffs, N.J.: Prentice Hall, 1958), p. 152.

predictive value."⁵ Thorndike and Hagen state that, ". . ., this performance inevitably depends in some measure upon the experiences that the individual has had."⁶

The foregoing does not essentially lead to the idea that an aptitude test must lay its emphasis entirely on the individual's inherited abilities nor does it imply that in the measurement of aptitude, the tester should guard against the element of the examinee's past achievement.

Generally, it is very difficult to prevent the simultaneous measurement of prior achievement in aptitude tests, but if items with which examinees have equal familiarity are written, the test^{variance} will tend to give more of a measure of the examinee's capacity than of his mere achievement.^{6A} The present test consists of items that are based mainly on experiences out of school, available to all. Also, the present test is developed on the basis of the analysis of the graduate program in Education at A.U.B. Thus, it is hoped, the performance of individuals on this test will differentiate them according to individual differences in capacity to do graduate work.

Limitations of the Study

For the purposes of the present study it was decided that only objective-type items were to be constructed, so as to insure an estimate of each item's dependability. Thus some of the major abilities had to

⁵ Frank S. Freeman, Theory and Practice of Psychological Testing (New York: Holt, 1950), p. 263.

⁶ Robert L. Thorndike, and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, (New York: John Wiley, 1955), p.204.

^{6A} Always assuming that items averaging about .50 in difficulty are written.

be ruled out due to the difficulty of finding objective ways to measure them.

The small size of the sample of subjects for whom the test is designed posed another limitation to the study. The tryout sample is small, but it is fairly a representative one inasmuch as the present group appears, in the judgments of several independent Education faculty observers, to be closely similar to groups of graduate students found in the Education Department at other times.

Due to the small size of the sample it was planned to retain all items which, on analysis, produced a chi-value⁷ of .70 or higher. Such items have a better than 50 percent chance of proving to be significantly reliable when the results of future testings are available. In fact, a chi-value of nearly 2.00 is required if the discriminative power of an item is to be clearly established at the .05 level of significance.⁸ Eight items between the chi-values 0.2 - 0.7 were retained in order to increase the average difficulty of the test. Thus there are items included in the revised test which are not significantly discriminative at the usual levels, and therefore it is possible that in future testings they may fail to discriminate between good and poor performers on the test. Thus this is a third limitation of the study.

As factors relating to personality, and measures of intelligence etc., though important, are considered to be separate studies in themselves,

⁷For details, see pp. 31-33 of the thesis.

⁸The discriminative power of such items is so high that there are only five chances in a hundred that the apparent discrimination between high-scorers and low-scorers could occur merely through chance sampling factors.

they have not been investigated in the present study. It is possible that the proposed test, no matter how valid it may turn out to be, would operate more efficiently as one unit of a battery of tests which takes other factors into account.

CHAPTER TWO

METHOD OF STUDY

Introduction

The first step in the construction of a test is to have the objectives of the test clearly formulated. The nature of these objectives determines the type and kind of items to be written.

The second step involves defining the objectives in terms of specific behaviors students are expected to demonstrate. For example, an objective such as the ability to locate relevant information may be defined in terms of specific behavior as the ability to select relevant data from a printed matrix of relevant and irrelevant information.

The third step consists of providing test situations which will give evidence of the presence of the behavioral objectives desired.

Fourth, a fairly representative sample of subjects for whom the test is intended are presented these situations. Then the test's reliability (before revision) is determined. Items are analyzed and either rejected or edited, and the test is ready (if sufficiently reliable) for validation study.

This standard procedure in test construction has been followed in the present study. However, the present study does not aim at finding

the validity of the test, a step which would require more time than is available, and which it is hoped will be carried out independently.

This chapter includes the statement of the objectives of the test, and data pertaining to tryout studies. Reliability data are presented in the next chapter.

Definition of the Term Academic Potential

In any prediction study, the analysis of the job for which the test is constructed precedes the task of constructing the test. It is essential to define what the job is composed of, and to analyze each component in terms of abilities which the test is to measure. As the purpose of the present study is to construct a test to predict academic potential for graduate work in Education, it is desirable to analyze the term academic potential: to determine what aptitudes and significant abilities are predictors of success in graduate work in Education at the American University of Beirut. The following procedures were resorted to in analyzing the term -- academic potential.

Graduate students in the Department of Education were interviewed as to what particular difficulties they encounter in their work, and what aspects of graduate work are most emphasized by their teachers and the graduation requirements.

Professors of the Education Department who are well-acquainted with the Department's graduate work in its many phases, were interviewed as to what particular abilities they thought were required and were worth looking for, if one was to determine whether a particular candidate would succeed in graduate work.

The researcher's background and personal experience of graduate work in Education was the basic resource for interpreting the interviewees' reports and for compiling lists of relevant and significant abilities.

By the above means, the following significant abilities, knowledges and skills required for success in graduate work in Education at the American University of Beirut were isolated:

1. The ability to interpret data.
2. The ability to locate relevant information.
3. The ability to derive inferences.
4. Reasoning ability: The ability to think logically and deductively.
5. The ability to organize data.
6. Basic Language Skills.
7. The ability to comprehend relations:
 - Analogies
 - Physical relations
8. The ability to comprehend written material.
9. Knowledge of appropriate sources.
10. General information.

The foregoing analysis does not, however, include some abilities which are possibly quite relevant, such as the ability to produce clear written expression, the ability to adjust to new situations, and other factors relating to the personality of the graduate student; these were stressed by many interviewees. These might be of crucial importance but they have been left out in the present study for the following reasons: the ability to produce clear written expression could not be included in the test because its adequate testing requires essay-type questions and these were ruled out because of their non-objective nature. Factors

relating to personality could not be included due to the limited time at the researcher's disposal and the fact that personality tests, while important, are generally regarded as separate studies in themselves.

The analysis thus provided the general objectives which were then defined in terms of student behaviors (activities) the test is intended to measure.

Objectives of the Test in Behavioral Terms

Each of the abilities, knowledges and skills (p. 11) was translated into specific behaviors the test is intended to measure:

Abilities, Knowledges and Skills	Objectives of the Test in Behavioral Terms
1. The ability to interpret data	1. Reads and gets information from line, bar and circular graphs.
2. The ability to locate relevant information.	2. Selects relevant data from a printed matrix of relevant and irrelevant information.
3. The ability to derive inferences.	3. Recognizes the unsoundness of drawing inferences from insufficient data or evidence.
4. Reasoning ability: The ability to think logically.	4. Reasons logically and deductively and applies generalizations appropriately.

5. The ability to organize data.

6. Basic Language Skills.

7. The ability to comprehend relations;

- Analogies

- Physical relations

8. The ability to comprehend written material.

9. Knowledge of appropriate sources.

e.g.: by the use of syllogistic techniques to gain insights into thought relationships.

5. Fills a given skeleton outline with the appropriate title, to produce a well-organized essay-outline.

6.(i) Understands the meanings of some common words and terms used in graduate work.

(ii) Uses commas, periods, quotation marks etc., and uses capital letters at appropriate places in a given sentence.

7. Understands relationships between words and ideas, and grasps physical relations in simple scientific phenomena.

8. Understands written material, and identifies true and false statements about it.

9. Understands when each of the following is the most appropriate source:

10. General information.

1. the library card catalog
2. the Reader's Guide to Periodical Literature.
3. the Education Index
4. Book Review Digest, etc.

10. Understands, for example, that the I.C.A. is not a U.N. agency.

Construction and Selection of Items

Objectives of the test were grouped into selections and subsections to facilitate item-writing. It was decided that the test should consist of objective items of the true-false, multiple-choice and completion types. Objective tests were chosen because they have marked advantages over essay tests in that they make it possible to include a large number of items that better sample the behaviors to be measured, and they make scoring easy and objective. Well-constructed objective tests are also devoid of factors such as ambiguity and vagueness of questions, quality of handwriting in answers, etc.

The construction of test items proceeded through the following four stages:

First, the following tests making use of a section or a subsection of any of the desired objectives were studied:

Acorn National Aptitude Tests *

Concept Mastery Test **

Progressive Reading Tests***

About eight items from the above tests were modified to suit the objectives of the present study and were included in the tryout test.

Second, items were written separately for each objective. One hundred and five items were originally written with a view to having a few in excess. It was the aim to include some items that most subjects could do, and also to include some items which would be difficult enough to spread out the ablest examinees. It was also intended that most items should be of moderate difficulty for the group as a whole, since such items make the greatest number of individual discriminations.

Third, each of the objectives was weighted on the basis of rational judgment in percentage terms as shown in Table II.

*By Andrew Kobal, et al., (New York: Acorn Publishing Company).

**By Lewis M. Terman (New York: The Psychological Corporation, 1950).

***By California Test Bureau.

TABLE II

DISTRIBUTION OF WEIGHT OVER OBJECTIVES OF THE TEST

(Total time of tryout test = 60 minutes).

Percent of weight Desired	General Objectives of the Test.	Minutes available	Minutes available for answering one item.	No. of answers in each item.	Score-points in each sub-section	Number of items
15	I <u>Ability to locate and interpret relevant data</u>	9				
5	(i) Ability to locate	3	$1\frac{1}{2} - 2\frac{1}{2}$	5	5	1
10	(ii) Ability to interpret	6	$3 - 3\frac{1}{2}$	4	8	2
30	II <u>Ability of reasoning and logical thinking</u>	18				
12	(i) Reasoning ability	7	1	1	8	8
6	(ii) Ability to derive inferences	4	3 - 4	5	5	1
12	(iii) Ability to organize data	7	$3/4$	10	10	1
20	III <u>Basic Language Skills</u>	12				
8	(i) Vocabulary	5	$\frac{1}{2} - 1$	1	7	7
4	(ii) Punctuation	2	$\frac{1}{2}$		4	4
8	(iii) Correct usage	5	$\frac{1}{2} - 1$	1 - 2	6	4
10	IV <u>Comprehension of relations</u>	6				
6	(i) Analogies	4	$3/4$	1	5	5
4	(ii) Physical relations	2	1 - 2	1	2	2
15	V <u>Misc. Skills</u>	9				
12	(i) Comprehension of written material	7	3 - 4	5	10	2
3	(ii) Knowledge of appropriate sources	2	1	1	2	2
&						
10	VI <u>General Information</u>	6	$3/4$	1	9	9
					81	

The weight attached to each objective was determined on the basis of its relative importance as well as the time a graduate student, on the average, devotes to that particular activity or set of activities. The weighting of objectives was further accomplished by manipulating the time. It was decided that the pilot test should not take more than an hour to avoid scheduling difficulties etc. This time was proportionately distributed over all sections and, in turn, over sub-sections. The number of items that could comfortably be attempted by most examinees was then determined, considering the time-factor of secondary importance. The percent of weight allocated to each objective determined within fairly narrow limits the number of score-points each section would carry.

Fourth, the items were re-read in consultation with appropriate professors in order to remove any apparent ambiguities in them. In the light of the critiques received, many items were either eliminated or revised.

After the editing of items was done, those to be included in the tryout test were selected. Items with the same type of answer-pattern were grouped together so that repeated instructions could be held to a minimum.

For clarity and for the convenience of both the examinee and the scorer, items were so arranged that all the parts of an item would appear on the same page and that a simple key could be used conveniently.

Administration of the Tryout Test

For the purpose of the tryout, the portions of the general

instructions on the front page of the tryout test' - "It is therefore to your advantage to make a guess on an item you are not sure about. You may take as long as you like" - were emphasized and read out to each of the fifty subjects who took the tryout test.¹

Though separate instructions with illustrative items are given for each section in the tryout test, it was observed that quite a few subjects appeared to find the instructions of sections D and F ambiguous; at any rate they complained that they could not understand what the instructions actually required them to do. These subjects were directed to reread the instructions for these sections carefully and try to understand the nature of the example preceding the items in the section. The instructions for sections D and F of the tryout form were revised for the final test² ^{so} as to reduce their apparent ambiguity.

The Nature of the Sample and the Testing Procedure

The tryout sample comprised fifty subjects. Five of them were M.A. graduates of the Department of Education, American University of Beirut. Thirty-nine were graduate students who had been working from one to six semesters of graduate work in Education. Nine of these also had spent one or two ~~summers~~^{semesters} of graduate work in Education, and are now actively pursuing their M.A. studies. Five were special students who had had one or two regular semesters of graduate work in Education; they are also pursuing their studies for the M.A. degree

¹ A copy of the tryout test with item-analysis results is included in Appendix A.

² A copy of the Revised Form is included in Appendix B.

in Education. There was one senior student who is intending to begin graduate work after June, 1960.

The test was conducted in the presence of the experimenter. Generally, the test was completed in one sitting. About fifteen such sittings were held because all the subjects could not come at one time, due to schedule difficulties etc. Two subjects completed the test in two sittings. Whatever the subject did in the first sitting was scored separately. The examinee was told about this before the second sitting, which took place from six to eight hours later.

The time when the examinee started answering test items was noted. When he had finished the test and turned in his test copy, the time was again noted. The difference between the two readings gave the approximate time the student required to finish working on the tryout test.³ The average time required for the tryout test was 61.4 minutes, ranging from 33 to 95 minutes.

³ A frequency distribution of time required is given in Table III.

TABLE III
FREQUENCY DISTRIBUTION OF TIME REQUIRED FOR THE
TRYOUT TEST

Time in Minutes	Frequency
95 - 99	1
90 - 94	0
85 - 89	0
80 - 84	2
75 - 79	12
70 - 74	3
65 - 69	4
60 - 64	6
55 - 59	1
50 - 54	8
45 - 49	5
40 - 44	6
35 - 39	1
30 - 34	1
Total = 50	

Mean = 61.4

Scoring

To facilitate scoring, a scoring key was made for each of the test pages by encircling the correct response to each item on the key and cutting rectangular slots in the key in such a way that when it was matched with the appropriate answer sheets to be scored, the key would indicate the wrong and the correct responses and omissions on the answer sheet. The wrong responses and omissions, when located, were marked on the answer sheets, which were then counted and subtracted from the total available score-points (81) on the test to find the number of answers given correctly, which was used as the raw score. No scoring formula involving a correction for chance success was used, as this would have involved instructing examinees not to guess on doubtful items. It is in the best interests of a tryout test, however, to have as many answers given as possible.

There was no scoring key for section F (items 27 - 30) because these items involved inserting punctuation marks and removing or adding capital letters in sentences. In this case, where a required comma or capital letter was missed or wrongly placed the whole item was considered to be wrong.

In item 26 (section E), for each of the options A, B, C, or D which was encircled, one score-point was given, and one point more for failing to encircle E.

Items 66 - 75 (section K) were scored as follows: Full credit (ten score-points) was given when the entire skeleton outline was filled correctly; when only the title was incorrect (the headings etc.,

being correct), two score-points were deducted. If subheadings of a group of items were correct but the corresponding heading did not belong to that group, no credit was given for that group of items. One score-point was given for a group of items when the subheadings and the related headings of the group were simply put together, but were not arranged according to the key. For example, one score-point is given when III, IIIA, and IIIB in the skeleton outline are considered to be 68, 71 and 69 (all of these items belong to group III, but three score-points are given when 71 is definitely assigned to III, and 68 and 69 are shared by IIIA and IIIB in any arrangement).

A randomly selected set of eighteen test copies was given to another graduate student for scoring so that the reliability of scoring could be ascertained. The rescoring was done with the scoring keys with additional instructions for sections E, F, and K attached. No discrepancy was found between the two sets of scores.

After a tryout copy was scored, the sum of mistakes and omissions was written on the bottom right-hand corner of the test copy. The net raw score (mistakes and omissions subtracted from the total of 81 possible score-points on the test) was written at the top left-hand corner of the test copy. This was done for convenience in checking the scores at a later period. The distribution of scores on the tryout test is given in Table IV. The mean and standard deviation for the tryout test, as computed from Table IV, are 61.3 and 8.22 respectively.

TABLE IV
FREQUENCY DISTRIBUTION OF SCORES ON THE
TRYOUT TEST

Scores	Frequency
76 - 78	1
73 - 75	1
70 - 72	6
67 - 69	4
64 - 66	11
61 - 63	8
58 - 60	5
55 - 57	5
52 - 53	4
49 - 51	2
46 - 48	0
43 - 45	0
40 - 42	2
37 - 39	1
Total = 50	

Mean = 61.3

Standard Deviation = 8.22

CHAPTER THREE

ANALYSIS OF THE TRYOUT DATA

The Purpose of the Analysis

The analysis of the tryout data aims to serve the following purposes:

1. To identify overly easy, overly hard, negatively - discriminating, ambiguous, weak and non-functioning items.
2. To obtain data on the level of difficulty of each individual item so that items may be selected to suit the purpose of the final test.
3. To obtain an index of the discriminating power of each individual item, so as to select items most likely to render the final measuring instrument most efficient.
4. To obtain data to compute a coefficient of reliability of the test as a whole.

The Analysis-Procedure

For the purpose of the present analysis, test papers were arranged from high to low on total raw score and the top and bottom twenty-seven percent (fourteen cases) were tallied. The numbers of right answers, the numbers and kinds of wrong choices, and the numbers

of omissions in an item were recorded for both the top and bottom groups. This was done individually for each item, as illustrated in Table V.

TABLE V
 ITEM-ANALYSIS DATA FOR ITEMS
 1, 2, 10, 16, 33 and 62 IN
 THE TRYOUT TEST

Item	Choices	Top 27 percent	Bottom 27 percent	Remarks	Key
1	A	0	0	NFM	<u>Non-functioning mislead</u>
	B	0	0	NFM	"
	C	0	0	NFM	"
	Ⓓ	14	14	NDA	<u>Non-discriminating answer</u>
2	A	0	0	NFM	<u>Non-functioning mislead</u>
	B	0	0	NFM	"
	Ⓒ	13	14	nFA	<u>Negatively-functioning answer</u>
	D	1	0	nFM	<u>Negatively-functioning mislead</u>
10	A	2	3	FM	<u>Functioning mislead</u>
	B	1	4	FM	"
	C	1	0	nFM	<u>Negatively-functioning mislead</u>
	D	0	2	FM	<u>Functioning mislead</u>
	Ⓔ	10	5	FA	<u>Functioning answer</u>
16	her	0	3	FM	<u>Functioning mislead</u>
	Ⓢhe	14	11	FA	<u>Functioning answer</u>
33	Ⓙ	11	11	NDA	<u>Non-discriminating answer</u>
	F	3	3	NDM	<u>Non-discriminating mislead</u>
62	T	0	2	FM	<u>Functioning mislead</u>
	Ⓡ	13	5	FA	<u>Functioning answer</u>
	F	1	6	FM	Functioning mislead
	omit	0	1		

The encircled choice in Table V is the correct answer as given by the scoring key. The other choices are misleads designed to attract examinees to whom the correct response is not obvious.

After the data about an item was recorded as above, each of the choices in the item was marked FA (for functioning answer, if the number of correct answers in the top 27 percent was greater than that in the the bottom 27 percent); FM (for functioning mislead, if more persons in the bottom group chose it than in the top group); NDA (for non-discriminating answer, if the number of persons choosing it in the top group was not zero and was equal to that in the bottom group); NDM (for non-discriminating mislead, if the frequency of choosing it in the top group was not zero and was equal to that in the bottom group); NFA (for non-functioning answer, if no persons in either group chose it); NFM (for non-functioning mislead, if no persons in either group chose it); nFA (for negatively-functioning answer, if those choosing it in the top group were fewer than those in the bottom group); nFM (for negatively-functioning mislead, if those choosing it in the bottom group were fewer than those in the top group).

The data in Table V indicates that the answers to items 1 and 2 in the tryout test are non-discriminating and negatively-functioning ones respectively. Besides, almost all other options in these two items are non-functioning misleads. Hence, these two items were discarded, and therefore are not included in the revised test. These are given as examples of this type of item data. Twenty-four others behaved similarly and were discarded.

In item 10, except option C which is a negatively-functioning

mislead, all other options are functioning ones. It is of moderate difficulty, as 15 persons out of 28 have answered it correctly. It was edited by deleting option C, and included in the revised test as item 7. There were eleven other items of this type which were retained after minor revision or deletion of one or two options.

Though item 33 has a non-discriminating answer, it is retained in the revised test (as item 19) in the hope that it will discriminate between individuals if it is used with relatively large samples. There was one more item with non-discriminating answer thus retained.

Items 16 and 62 discriminate well because almost all persons in the top group have chosen the correct answer. In item 62, both misleads, T and F, have attracted eight persons out of fourteen in the bottom group. It is moderately hard as only 18 persons out of 28 have answered it correctly. Both items, 16 and 62, are retained (as items 9 and 38 respectively) in the revised test. Thirty-seven such items which make forty-one score-points, were retained without change.

The above data about each item helped in understanding the item as regards its level of difficulty, amount of discriminating power, usefulness, weaknesses, etc.¹

Odd versus even items

The whole (tryout) test was divided into two presumably equivalent halves. For the purpose of simplicity and objectivity, alternate items were put into the two half-tests, that is, all odd-

¹See also Selection and Revision of Items (p.36) for further details on item selection and revision.

numbered items in the tryout were considered as subtest X, and all even-numbered items as subtest Y, except for a few items that required two answers; in these cases the two answers were alternately put into the two subtests. The mistakes and omissions of every subject in subtests X and Y were tallied separately. Thereby, each subject obtained an odd and an even score. The frequency distribution of these scores in subtests X and Y, given in Table VI, provides means of 30.6 and 30.7 respectively, and standard deviations of 4.20 and 3.94 respectively. Tests for differences between the two means and the two standard deviations showed no significant differences.

This splitting of the test was done so as to obtain an independent reliability estimate as a check on the Kuder-Richardson estimates. Reliability measures are discussed on pages 33 to 36 inclusive.

TABLE VI
DISTRIBUTION OF SCORES ON SUBTESTS X AND Y OF
THE TRYOUT TEST

X - Scores		Y - Scores	
Score	Frequency	Score	Frequency
40 - 41	1	40 - 41	0
38 - 39	0	38 - 39	1
36 - 37	6	36 - 37	1
34 - 35	5	34 - 35	11
32 - 33	8	32 - 33	15
30 - 31	10	30 - 31	6
28 - 29	10	28 - 29	7
26 - 27	7	26 - 27	3
24 - 25	0	24 - 25	2
22 - 23	1	22 - 23	0
20 - 21	1	20 - 21	4
N = 50		N = 50	
Mean _(X) = 30.6		Mean _(Y) = 30.7	
Standard deviation = 4.20		Standard deviation = 3.94	

Item Difficulty

The difficulty level of each individual item was expressed in terms of the ratio of right answers in the top and bottom twenty-seven percent of the sample to the total number of cases in these two groups. For example, the level of difficulty (D) of item 62 (Table V) was calculated as

$$\frac{13 + 5}{28} = .64$$

The data in Table VII gives measures of difficulty (D) and significance of discriminative power (Chi) of individual test items. The average item difficulty of the tryout test is calculated to be .74.

Discriminative Power of the Items

The amount of discriminative significance of each item was calculated by Cureton's Chi test² in terms of chi-value, as given in Table VII. The value of chi for each item computed by the chi-test was compared with the Cureton's Table³ of Chi-values at various Significance Levels for certain Sample Sizes, to see if chance could account for the item's apparent discriminative power.

²Frederick B. Davis, "Item Selection Techniques," Educational Measurement, ed. E.F. Lindquist (Washington D.C.: American Council on Education, 1951), p. 289.

³Ibid., p. 290.

TABLE VII
 MEASURES OF LEVEL OF DIFFICULTY (D) AND DISCRIMINATING
 SIGNIFICANCE (CHI) OF INDIVIDUAL ITEMS IN
 THE TRYOUT TEST

Item	D	Chi	Item	D	Chi	Item	D	Chi	Item	D	Chi
1	1.00	0	23	.86	.546	43	.89	0	67	.82	1.97
2	.96	0	24	.79	1.4	44	.93	.74	68	.86	1.77
3	1.00	0	25	.93	.74	45	.93	.74	69	.72	2.96
4	.86	.546	26A	.82	1.97	46	.61	1.55	70	.72	2.96
5	.96	0	B	.68	1.68	47	.46	.757	71	.68	3.2
6	.61	1.55	C	.50	1.89	48	.86	.546	72	.86	1.77
7	.86	1.64	D	.82	1.97	49	.68	0	73	.68	2.42
8	1.00	0	E	1.00	0	50	.89	0	74	.72	2.96
9	.89	1.20	27	.04	0	51	.72	1.27	75	.79	2.32
10	.54	1.52	28	.04	0	52	.04	0			
11	.57	.382	29	.54	1.52	53	.82	0			
12	.72	.423	30	.50	1.89	54	.57	4.2			
13	.36	.396	31	1.00	0	55	.54	1.52			
14	.79	.465	32	.82	.982	56	.96	0			
15	1.00	0	33	.79	.468	57	.28	1.25			
16	.89	1.20	34	.89	0	58	.86	.546			
17(I)	.68	1.68	35	.82	.985	59	.86	1.64			
17(i)	.68	1.68	36	1.00	0	60	.36	1.19			
18	1.00	0	37	.86	1.64	61	.54	3.04			
19	1.00	0	38	.89	1.20	62	.64	2.76			
19	.75	0	39	.64	1.97	63	.64	2.76			
20	.68	0	40	.68	1.62	64	.72	2.96			
21	.72	1.27	41	.86	.546	65	.75	1.75			
22	1.00	0	42	.89	0	66	.72	2.96			

It was found that twelve items were discriminative at the .05 level of significance; out of these, ten items were discriminative at the .01 level of significance. The top and bottom twenty-seven percent of the sample-group were utilized to compute the values of chi. According to F.B. Davis, the Chi test ". . . is applicable to a greater degree than most others when the samples are small . . ." ⁴ to determine the significance of item discrimination.

Reliability of the Test

Reliability is one of the essential characteristics of a test. The reliability data give an indication of the extent to which the test measures accurately and consistently whatever it does measure. To the extent that the results of a measuring instrument are reproduced on repeated measurements of the same quantity, it is reliable. For example, if a test is administered twice to the same examinees (who have undergone no systematic changes), and the scores of the first time correlated with those of the second time, the resulting coefficient of correlation will be the test-retest index of reliability of the test.

The first step after the mean and standard deviation of the tryout test (Table IV) was obtained was to determine the reliability of the test by means of the Kuder-Richardson Formula No. 21. ⁵

$$r_{tt} = \frac{n}{n-1} \left[1 - \frac{M_t \left(1 - \frac{M_t}{n}\right)}{s_t^2} \right], \text{ where } r_{tt} \text{ is the}$$

⁴For details see Ibid., p. 288.

⁵Robert L. Thorndike, and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, (New York: Wiley & Sons, 1955). p. 131.

estimate of reliability, n is the number of items in the test, s_t is the standard deviation of the test, M_t is the mean score of the group. This formula gave a lower-limit estimate of +.79 as the test's reliability.

Formula No. 21 is generally used only as a rough approximation. It assumes equal item difficulty throughout. Any differences in item difficulty which exist may be expected to raise the reliability, and thus Formula No. 21 is generally understood to be a lower-limit estimate. A better estimate of the test's reliability was made by using Kuder Richardson Formula No. 20.⁶

$$r_{tt} = \frac{n}{n-1} \left\{ \frac{s_t^2 - \sum p_i q_i}{s_t^2} \right\}, \text{ where } r_{tt} \text{ is reliability of}$$

the total test, n is number of items in the test, s_t^2 is variance of the total test, \sum means "the sum of", p_i = proportion passing item i (=D), $q_i = 1-p_i$ (=1-D). The test's reliability coefficient obtained by the above formula is +.835. This formula is more likely to estimate the true reliability, as it takes item difficulty into account.

The split-half method was also applied as a third check on the reliability of the test. Retesting on the same test could not be conducted for administrative reasons. When the scores of individuals on subtest X (odd items) were correlated⁷ with their scores on subtest Y (even items), a coefficient of correlation of +.74 was obtained.

⁶Robert L. Thorndike, "Reliability," Educational Measurement, ed. E.F. Lindquist (Washington D.C.: American Council on Education, 1951), p. 587.

⁷Diagram 1 is the scatter diagram of correlation between subtests X and Y (odd and even items respectively).

The coefficient of correlation, thus obtained, was adjusted for double length by the use of Spearman-Brown Correction formula.⁸

$$r_{tt} = \frac{2r_{hh}}{1 + r_{hh}}, \text{ where } r_{tt} \text{ is the coefficient of}$$

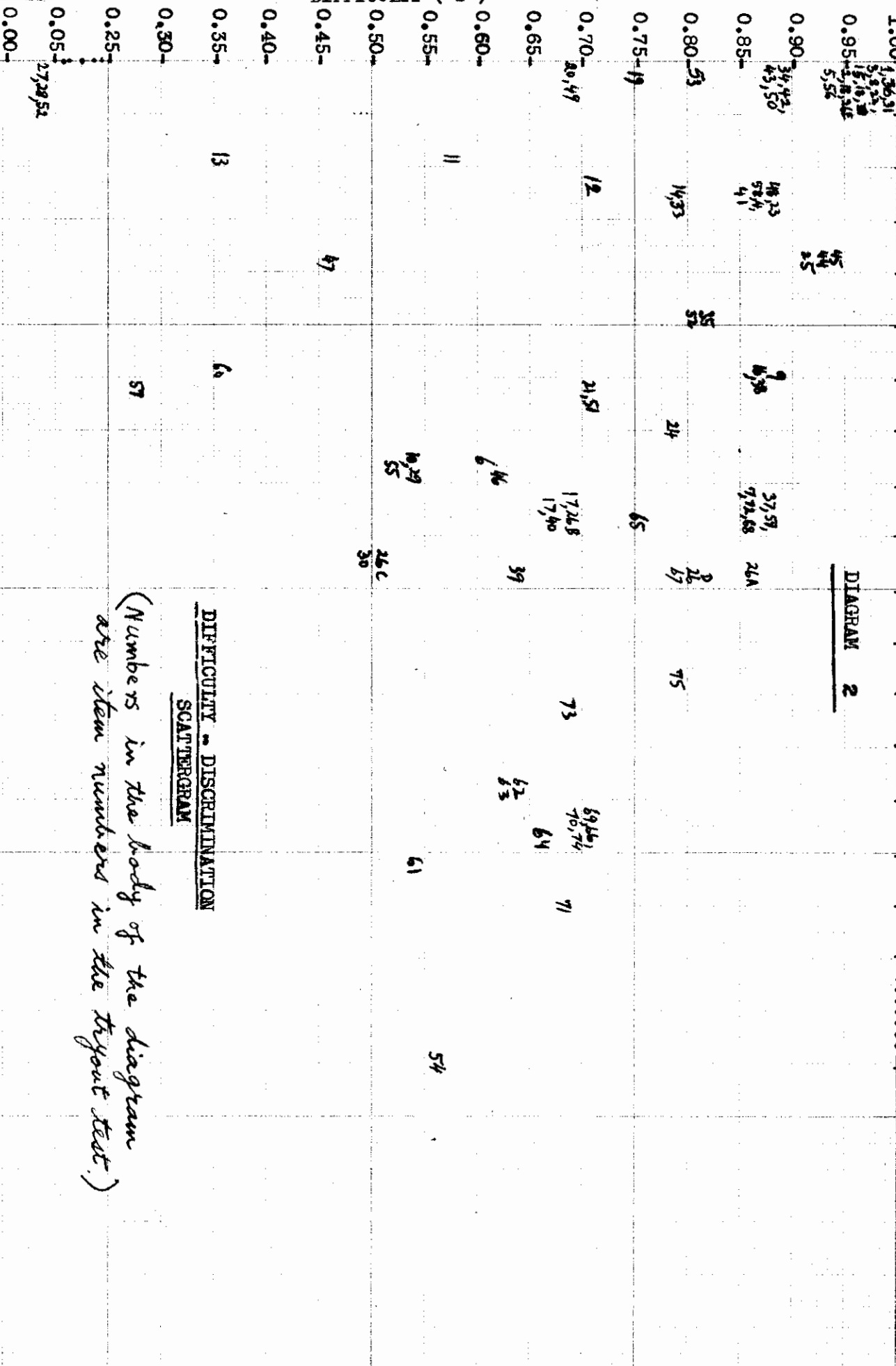
reliability for the total test, and r_{hh} is the coefficient of correlation between the scores on the two half-tests (X and Y) of the test. Using the above formula, a coefficient of reliability of +.851 for the total test was obtained, which agrees well with the Kuder-Richardson estimate. Elimination of poor items may be expected to raise this value appreciably.

Selection and Revision of Test Items

The level of difficulty of each item was plotted⁹ against the chi-value of that item, giving a difficulty-discrimination scattergram. This was done for all items in the test, as given in diagram 2.

⁸ J.P. Guilford, Fundamental Statistics in Psychology and Education, (New York: McGraw, 1950), p. 492.

⁹ This method of item-evaluation was suggested by Professor Korf.



DIFFICULTY - DISCRIMINATION
SCATTERGRAM

(Numbers in the body of the diagram are item numbers in the layout test.)

It can be seen from the diagram that most of the items are above the .60 level of difficulty, showing thereby that, on the whole the test tended to be on the easy side for the examinees.

In order to make the present instrument more efficient, and to select items for the final test, it was planned that, with only a few exceptions to be mentioned later, all items below the chi-value or 0.70 in diagram 2, should be discarded.¹⁰ Thus twenty-six items were dropped. All items that were at or above the chi-value 0.70 in the diagram were retained except item number 25, due to its clearly ambiguous nature.¹¹ Of those items falling between the chi-values of 0.2 and 0.70, it was decided to retain those items that fell below D-value 0.80. This was done to increase the average difficulty of the test. This gives the revised test a far higher probability of making a large number of discriminations between examinees.¹²

All of the preceding information about items was supplemented by the tabulated data for each item (Table V). In some items, one or two non-functioning or non-discriminating misleads had to be deleted, with the hope of improving the discriminative power of these items.

¹⁰The value 0.70 was chosen because any item having a chi-value higher than 0.70 is more likely than not to be a discriminating item when more subjects are tried on it. In exact statistical terms, the value 0.675 is the 50th Percentile of the chi distribution for one degree of freedom.

¹¹It was observed after the tryout, that the word 'paw' in the item may logically be considered as either a verb or a noun, and therefore two answers are equally reasonable.

¹²Statistical information supplied by Professor Korf, the advisor.

A few items were edited with such minor changes as would not necessitate a second tryout, with the hope that the change would increase the discriminative power of the item.

On examining choice-by-choice data for each individual item (Table V), seven edited items with a non-discriminating mislead in each, and three items that were above D-value .80 and between Chi-values of 0.20 and 0.70 were retained in the hope that with relatively large samples, they might discriminate between individuals in the desired direction. Out of the twenty-six items that were below the chi-value of 0.70 and were rejected, four showed negative discrimination (contained negatively-discriminating answers), and the rest were mostly non-discriminating, non-functioning, too easy, or too hard items.

The Revised Test

The revised test contains fifty items which make fifty-four score-points. For the ease of examinees in following different instructions for different types of tasks in the test, the original system of sections has been retained. The items have been placed in the order of their increasing difficulty in each section of the revised form. The instructions for sections D and F have been revised. The fact that many items included in the revised test are edited either by deleting one of their choices or replacing a choice by a presumably better one, suggests that on the average, the items in the revised test are more difficult than those in the tryout form. The average item difficulty of the revised test is estimated to be about .70, but of course the actual figure cannot be established until the revised items have been tried on actual examinees.

Conclusions and Recommendations

The frequency distribution of scores on the total test indicates that the distribution is negatively skewed, that is to say there are more high scorers than low scorers. The removal of many very easy items will balance the test more appropriately as far as the average item difficulty is concerned.¹³

The removal of some items and the editing of others on the basis of statistical analysis makes it probable that the reliability of the revised test will be higher than that of the tryout form (which showed a Kuder-Richardson Formula No. 20 reliability coefficient of +.835).

The statistical analysis of tryout results shows that there are at least ten items located in sections J and K of the tryout test (i.e., sections I and J of the revised form), which discriminated at the .01 level of significance. If these sections continue to be so discriminative and if high scorers on these two sections also succeed in graduate work, these two sections might serve better as eventual predictors of academic potential than the test as a whole. Only the eventual validity study could establish this, and the literature suggests strongly that such an outcome would be highly improbable, inasmuch as a short test generally lacks the reliability of a longer

¹³ According to Davis, the number of discriminations an item can make approach "its maximum where 50 percent of the examinees know the answer."

For details, see F.B. Davis, "Item Selection Techniques," Educational Measurement, ed., E.F. Lindquist (Washington D.C.: American Council on Education, 1951), pp. 308-309.

one, and because an item with a high item-test intercorrelation does not necessarily show a high correlation with a validity criterion.

Before the proposed test is actually used for prediction purposes, a validity study is necessary to determine whether high scorers on the test also succeed in graduate work.

The test clearly has "face validity" - i.e., in the judgments of experts the items appear to be measures of the aptitude in question. However, "face validity" is not necessarily true validity, which can only be established against a carefully worked out, objective criterion of success in graduate work. Even if the test is found to be reasonably valid, it would probably operate more efficiently as one unit of a battery of tests which takes factors relating to personality, measures of intelligence, etc. into account.¹⁴

It is hoped that the present study will serve as a starting point for future studies related to the graduate program in education at the American University of Beirut. It is also hoped that the present study will serve as an instrument which will help the Department of Education select potentially successful candidates for graduate work. This will not only save the time and resources of the Department, but will also help to raise the standard of graduate work in the Department

¹⁴ According to Mosier, if tests are combined to constitute a battery, e.g., an admissions battery for college entrance, their results "... are likely, in the light of much previous research, to give a better prognosis of college achievement than is any one test alone."

For details, see Charles I. Mosier, "Batteries and Profiles," Educational Measurement, ed. E.F. Lindquist (Washington D.C.: American Council on Education, 1951), p. 765.

by holding the number of failures to a minimum. Not least will be its usefulness in preventing some of the heartbreaks that so often accompany unsuccessful performance on the graduate level.

To sum up, the present study has succeeded in making items that appear to measure what the test is designed to do, and that have reasonably high reliability. It is recommended that the essential step of a validity study be conducted as soon as possible so as to establish the true validity of the test against a true criterion of success in graduate work.

APPENDIX A

The Tryout Test with Item-Analysis Results.

(The number in parenthesis for each item represents
the number of that item in the revised test).

EXPERIMENTAL
GRADUATE APTITUDE TEST

April, 1960

Department of Education, American University of Beirut

Name: _____
(Family name) (First) (Second)

Date: _____

Note to the Student About This Test:

This test is designed to measure the aptitude of future graduate students.

Directions: You will find different directions for different sections in the test. Please read these directions carefully.

Your score on this test will be the total number of items correctly answered. It is therefore to your advantage to make a guess on an item you are not sure about. You may take as long as you like, but please keep working so that an accurate estimate of the time for administration of the test may be made.

SECTION A

Directions: ENCIRCLE THE LETTER OF THE WORD OR PHRASE THAT BEST COMPLETES EACH OF THE FOLLOWING:

Example

To be tidy means to be

- A transparent.
- B neat.
- C dirty.
- D rough.

-
1. To comprehend means to
- (Discarded)
Chi = 0
D = 1.00
- A memorize.
 - B read.
 - C summarize.
 - D understand.
2. Adaptation means
- (Discarded)
Chi = 0
D = .96
- A accomplishment.
 - B to pursue one's job.
 - C adjustment.
 - D acceptance.
3. To be conscious means to be
- (Discarded)
Chi = 0
D = 1.00
- A hasty.
 - B devoted to duty.
 - C loyal.
 - D aware
 - E proud.
4. Adolescence refers to the period between
- (Discarded)
Chi = .546
D = .86
- A puberty and adulthood.
 - B childhood and puberty
 - C infancy and childhood.
 - D manhood and old age.
5. To compensate means to
- (Discarded)
Chi = 0
D = .96
- A concentrate.
 - B counterbalance.
 - C complain.
 - D complicate.
 - E to be able to compete.
6. To recapitulate means to
- (Retained
with editing -5)
Chi = 1.55
D = .61
- A receive.
 - B summarize.
 - C seize.
 - D catch hold of.
7. To be tentative means to be
- (Retained
with editing - 2)
Chi = 1.64
D = .86
- A taken for granted.
 - B conclusive.
 - C provisional.
 - D authentic.
 - E confidential.
8. Newton, Galileo and Archimedes are known chiefly as
- (Discarded)
Chi = 0
D = 1.60
- A poets.
 - B chemists.
 - C physicists.
 - D historians.
 - E anthropologists.

9. The embryo in the case of human beings is formed from (Retained with editing - 1)
- A the egg or the ovum of the mother. Chi=1.20
D = .89
 - B the sperm of the father.
 - C both the egg and the sperm.
 - D neither the egg nor the sperm.
10. A pair of identical twins is formed as a result of the (Retained with editing-7) Chi=1.52; D= .54
- A fertilization of one egg by two sperms.
 - B fertilization of two eggs by one sperm.
 - C fertilization of two eggs by two sperms.
 - D division of a sperm or an egg before fertilization.
 - E division of one egg after fertilization.
11. In weather forecasting the most important factor a meteorologist has to consider is change in the air's (Retained with editing - 6)
- A velocity. Chi = .382
 - B humidity D = .57
 - C pressure.
 - D temperature.
12. Taxes are charged mainly to (Retained with editing -4)
- A provide revenue for the government. Chi=.42
D =.72
 - B check inflation.
 - C to employ people as tax collectors.
 - D guard against accumulation of money in the hands of a few persons.
13. Your teacher advises you to quote a particular author's magazine article in your term paper. The best way to find his article will be to look in the (Retained with editing - 8)
- A Book Review Digest. Chi = .396
 - B Education Index. D = .36
 - C Encyclopedia Britannica.
 - D library's main card-index under the name of the article.
14. Suppose you are asked to locate a particular article dealing with school curricula published in a magazine. The easiest and the quickest way to do this is to look for the article in the (Retained with editing -3)
- A past issues of journals. Chi=.465
 - B Reader's Guide to Periodical Literature. D=.79
 - C Book Review Digest.
 - D library's main card-index under the name of the article.

9. The embryo in the case of human beings is formed from
- (Retained with editing - 1)
Chi = 1.20
D = .80
- A the egg or the ovum.
 - B the sperm of the father.
 - C both the egg and the sperm.
 - D neither the egg nor the sperm.
10. A pair of identical twins is formed as a result of the
- (Retained with editing - 7)
Chi = .82
D = .64
- A fertilization of one egg by two sperms.
 - B fertilization of two eggs by one sperm.
 - C fertilization of two eggs by two sperms.
 - D division of a sperm or an egg before fertilization.
 - E division of one egg after fertilization.
11. In weather forecasting the most important factor a meteorologist has to consider is change in the air's
- (Retained with editing - 8)
Chi = .508
D = .87
- A velocity.
 - B humidity.
 - C pressure.
 - D temperature.
12. Taxes are charged mainly to
- (Retained with editing - 4)
Chi = .488
D = .72
- A provide revenue to the government.
 - B check inflation.
 - C to employ people as tax collectors.
 - D guard against accumulation of money in the hands of a few persons.
13. Your teacher advises you to quote a particular author's magazine article in your term paper. The best way to find his article will be to look in the
- (Retained with editing - 8)
Chi = .508
D = .88
- A Book Review Digest.
 - B Education Index.
 - C Encyclopedia Britannica.
 - D library's main card-index under the name of the article.
14. Suppose you are asked to locate a particular article dealing with school curricula published in a magazine. The easiest and the quickest way to do this is to look for the article in the
- (Retained with editing - 5)
Chi = .488
D = .79
- A past issues of journals.
 - B Reader's Guide to Periodical Literature.
 - C Book Review Digest.
 - D library's main card-index under the name of the article.

9. The embryo in the case of human beings is formed from

- (Retained with editing - 1)
Chi = 1.20
D = .80
- A the egg or the ovum of the mother.
 - B the sperm of the father.
 - C both the egg and the sperm.
 - D neither the egg nor the sperm.

10. A pair of identical twins is formed as a result of the

- (Retained with editing - 7)
Chi = .52
D = .54
- A fertilization of one egg by two sperms.
 - B fertilization of two eggs by one sperm.
 - C fertilization of two eggs by two sperms.
 - D division of a sperm or an egg before fertilization.
 - E division of one egg after fertilization.

11. In weather forecasting the most important factor a meteorologist has to consider is change in the air's

- (Retained with editing - 8)
Chi = .502
D = .87
- A velocity.
 - B humidity.
 - C pressure.
 - D temperature.

12. Taxes are charged mainly to

- (Retained with editing - 4)
Chi = .455
D = .72
- A provide revenue to the government.
 - B check inflation.
 - C to employ people as tax collectors.
 - D guard against accumulation of money in the hands of a few persons.

13. Your teacher advises you to quote a particular author's magazine article in your term paper. The best way to find his article will be to look in the

- (Retained with editing - 8)
Chi = .308
D = .88
- A Book Review Digest.
 - B Education Index.
 - C Encyclopedia Britannica.
 - D library's main card-index under the name of the article.

14. Suppose you are asked to locate a particular article dealing with school curricula published in a magazine. The easiest and the quickest way to do this is to look for the article in the

- (Retained with editing - 5)
Chi = .455
D = .79
- A past issues of journals.
 - B Reader's Guide to Periodical Literature.
 - C Book Review Digest.
 - D library's main card-index under the name of the article.

SECTION B

Directions: ENCIRCLE THE WORD IN THE PARENTHESES THAT BEST COMPLETES EACH OF THE FOLLOWING SENTENCES:

Example:

He came to see (I)
(me)

BEGIN HERE:

15. You and I (are)
(am) leaving for Washington, tomorrow. (Discarded)
Chi = 0
D = 1.00
16. It was (her)
(she) that went out just now, wasn't it ? (Retained - 9)
Chi = 1.20
D = .89
17. (Its)
(It's) too bad that your dog has hurt (its)
(it's) leg. (Retained - 10)
Chi = 1.68
D = .68
18. Have you (wrote)
(written) to everyone who (wrote)
(written) to you ? (Discarded)
Chi = 0
D = 1.00

SECTION C

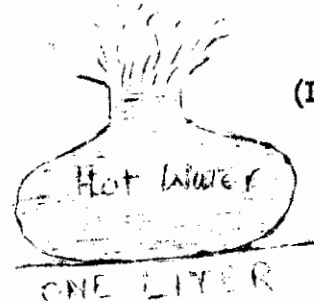
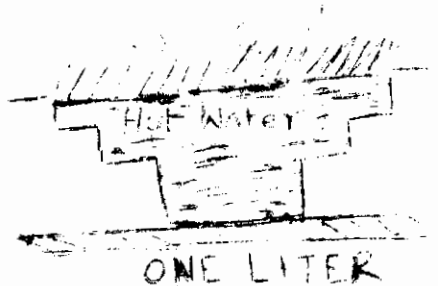
Directions: ENCIRCLE THE LETTER OF THE STATEMENT THAT BEST ANSWERS EACH OF THE FOLLOWING QUESTIONS:

19. A car ran 4 miles in the direction of the North and then 3 miles westwards. How long is the straight line between the car and its starting point.

- A 3 miles
- B 5 miles
- C 7 miles
- D 9 miles

(Discarded)
Chi = 0
D = .75

20.



(Discarded)
Chi = 0
D = .68

INFORMATION: Hot water gives heat to its surroundings in the process of cooling. Heat is given out more rapidly when the surface area exposed to the air is increased.

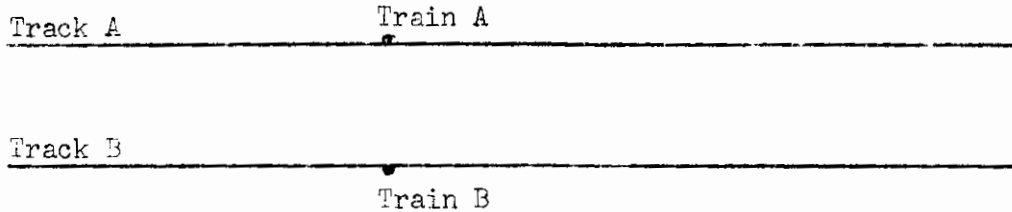
Which container will remain hot for a longer time ?

- A container A
- B container B
- C neither one; both of them will cool equally quickly.

SECTION E

Read the following passage carefully:

26. Two trains, A and B, start from the same place on parallel straight tracks at 10 A.M. It is desired to know how far apart they will be when the sun sets. (The distance between the tracks may be ignored.)



Directions: ENCIRCLE THE LETTERS OF THE FACTORS BELOW WHICH YOU THINK ARE ESSENTIAL TO THE SOLUTION OF THE ABOVE PROBLEM:

- (Retained - 14A) A the average speed of A. Chi = 1.97, D = .82
- (Retained - 14C) B the time of sunset. Chi = 1.68, D = .68
- (Retained - 14D) C the directions of movement of the two trains. Chi = 1.89, D = .50
- (Retained - 14B) D the average speed of B. Chi = 1.97, D = .82
- (Discarded) E the time when the sun is directly overhead. Chi = 0, D = 1.00

SECTION F

Directions INSERT PUNCTUATION MARKS AND ADD OR REMOVE CAPITAL LETTERS WHERE NECESSARY IN EACH OF THE FOLLOWING SENTENCES:

27. Sir said Jim I have seen these Instruments in the Laboratory but I

(Discarded)

don't know their use.

Chi = 0

D = .04

28. I wanted to wear it the first time tonight she declared when it's old

(Discarded)

I shall keep it and remember.

Chi = 0

D = .04

29. He washed examined his somewhat too easily growing beard and decided

(Retained - 15)

to shave again.

Chi = 1.52

D = .54

30. An exhibit will begin on monday afternoon at half past one everybody

(Retained - 16)

is invited the exhibit is free.

Chi = 1.89

D = .50

SECTION G

Directions: ENCIRCLE T IF THE STATEMENT IS TRUE OR F IF IT IS FALSE.

Example:

(T) F Grass is green.

BEGIN HERE:

31. T F Genes are the carriers of heredity. (Discarded)
Chi = 0, D = 1.00
32. T F I.C.A. is a U.N. agency. (Retained - 17)
Chi = .982, D = .82
33. T F Chronology is the science dealing with the measurement of time. (Retained - 19)
Chi = .468, D = .79
34. T F Latin America is to the north of the United States. (Discarded)
Chi = 0, D = .89
35. T F Red China is another name for Nationalist China. (Retained - 18)
Chi = .985, D = .82

SECTION H

Directions: In each of the following items, the third statement may be true or false, depending upon the two statements preceding it, which you must assume to be true.

ENCIRCLE T IF THE THIRD STATEMENT IS TRUE OR F IF IT IS FALSE:

Example:

All men are mortal
He is a man
(T) F Therefore he is mortal

BEGIN HERE:

- | | | | |
|-----|-----|---|--|
| 36. | | a's are equal to b's
b's are equal to c's
Therefore a's are equal to c's | (Discarded)
Chi = 0
D = 1.00 |
| | T F | | |
| 37. | | Some men live to eat
He is a man
Therefore he lives to eat | (Retained - 21)
Chi = 1.64
D = .86 |
| | T F | | |
| 38. | | All fruits have skins
Nuts are fruits
Therefore all nuts have skins | (Retained - 20)
Chi = 1.20
D = .89 |
| | T F | | |
| 39. | | All ploos nepp smoothly
All smooth neppers are ponns
Therefore all ploos are ponns. | (Retained - 24)
Chi = 1.97
D = .64 |
| | T F | | |
| 40. | | All plants contain chlorophyll
Mushrooms do not contain chlorophyll
Therefore mushrooms are not plants | (Retained - 23)
Chi = 1.62
D = .88 |
| | T F | | |
| 41. | | All toys are made of wax
This object is made of wax
Therefore this object is a toy | (Retained - 22)
Chi = .546
D = .86 |
| | T F | | |
| 42. | | Some swimmers are wrestlers
All wrestlers are tennis players
Therefore some swimmers are tennis players | (Discarded)
Chi = 0
D = .89 |
| | T F | | |

SECTION I

Directions: Read the following passage carefully. In front of the statements below it ENCIRCLE:

YES, if the statement is supported by the data in the passage,
NO, if the statement contradicts the data in the passage,
?, if the passage gives no evidence of the statement's being true or untrue.

(Adapted from TIME, Atlantic Edition, November 16, 1959, p. 55)

It will take a generation or more to clear the state mental hospitals of the backlog of patients permanently crippled by old-time procedures that, far from making them better helped to make them worse. But seclusion rooms are being converted into kitchenettes and beauty parlors; camisoles and straps are disappearing. Shock treatment is seldom used, and only for selected patients. Though admission rates are rising, release rates are rising faster, so that in many states there is a net decrease in the numbers of mentally ill confined to hospitals.

43. YES ? NO Old-time procedures made patients better.
(Discarded) Chi = 0, D = .89
44. YES ? NO Shock therapy is now used for some patients.
(Retained - 25) Chi = .74, D = .9
45. YES ? NO Admission rates are rising faster than release rates.
(Retained - 26) Chi = .74, D = .9
46. YES ? NO Throughout the country the proportion of mentally ill persons outside hospitals is increasing. (Retained - 27)
Chi = 1.55, D = .61
47. YES ? NO "make them worse" (line 3 of the passage) refers to the physical condition of patients. (Retained - 28)
Chi = .757, D = .46

Directions: Read the following passage carefully. In front of the statements below it ENCIRCLE:

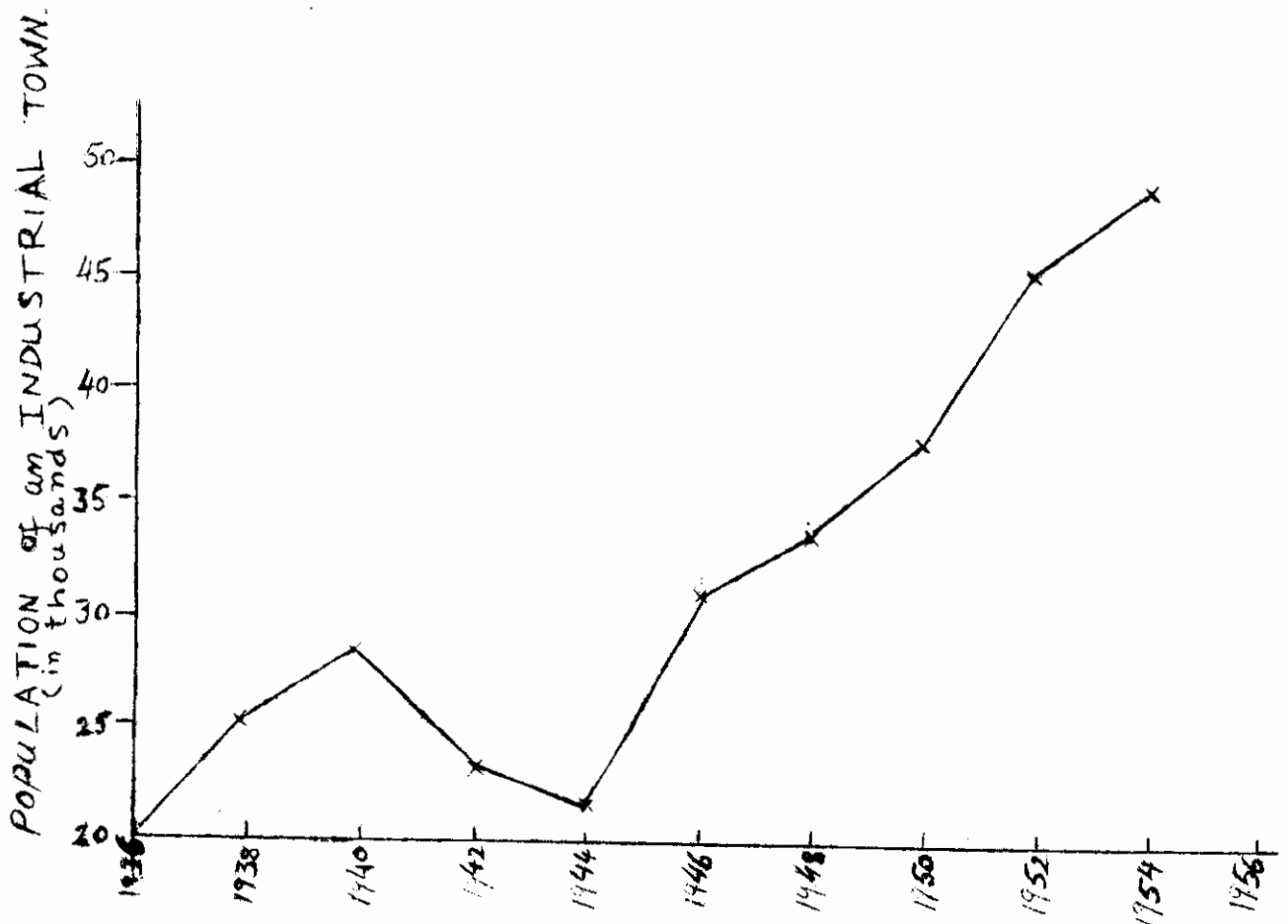
YES, if the statement is supported by the data in the passage,
NO, if the statement contradicts the data in the passage,
?, if the passage gives no evidence of the statement's being true or untrue.

(Adapted from Jones and Darkenwald, Economic Geography.)

The rate of population growth has not been uniform in all parts of the world. In some old lands population has declined, whereas in others, for example China and India, numbers have continued to multiply, reducing the standard of living to the subsistence level or a little above it. The population of France has been practically static for some time, even though there has been a marked change in the economic activities of the people. In the United Kingdom the population is gradually becoming stationary. In the United States the rate of increase is declining.

48. YES ? NO The population has increased by approximately equal amounts throughout the world. (Retained - 29)
Chi = .546, D = .86
49. YES ? NO The population has been decreasing in the United Kingdom. (Discarded)
Chi = 0, D = .68
50. YES ? NO For all practical purposes, there has been an increase in the population of France. (Discarded)
Chi = 0, D = .89
51. YES ? NO In Pakistan, the standard of living is near the subsistence level due to overpopulation. (Retained - 30)
Chi = 1.27, D = .72
52. YES ? NO In the United States the population is decreasing. (Discarded)
Chi = 0, D = .04

Directions: Study the following graph carefully. In front of the statements below it ENCIRCLE:
 YES, if the statement is supported by the data in the graph,
 NO, if the statement contradicts the data in the graph,
 ?, if the graph gives no evidence of the statement's being true or untrue.

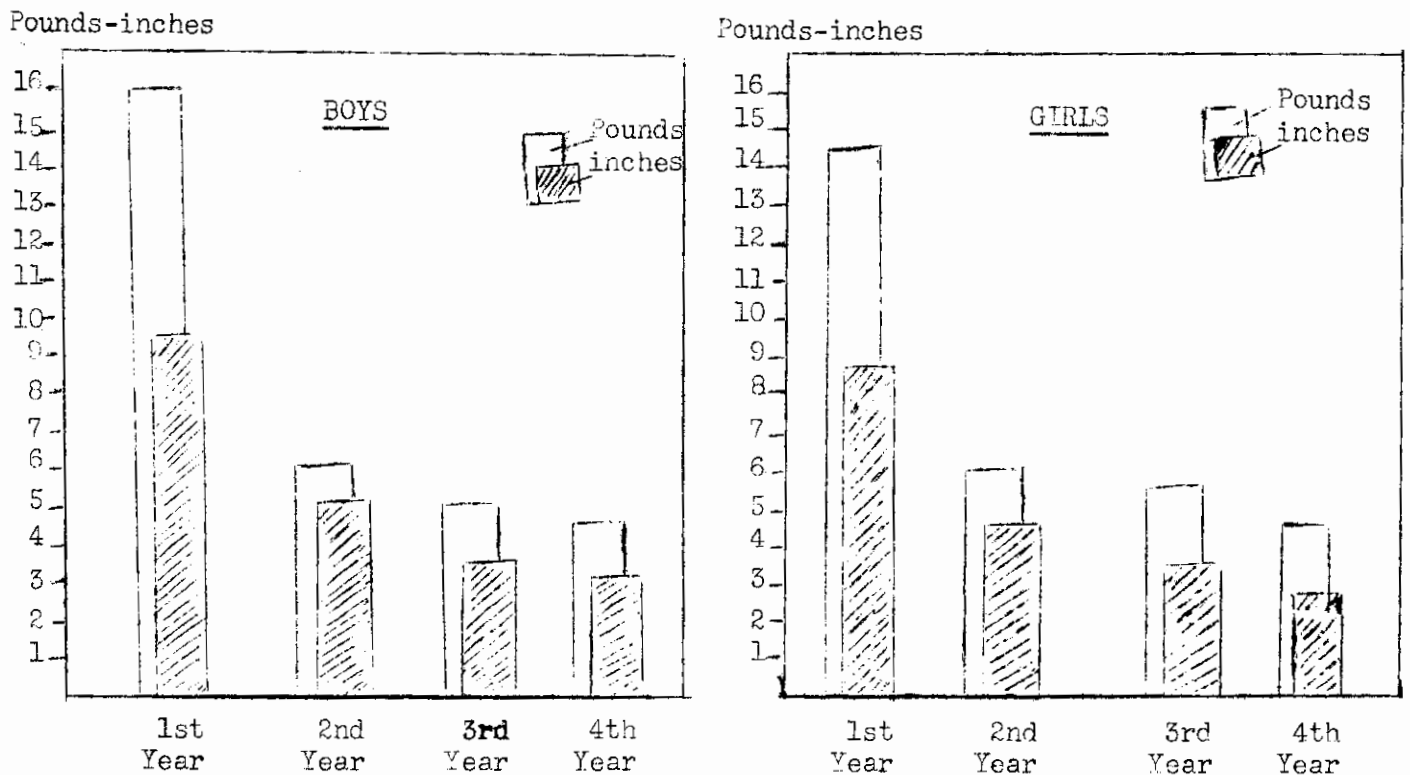


53. YES ? NO The greatest increase in population occurred during the period from 1950-52. (Discarded)
 Chi = 0, D = .82
54. YES ? NO Many people died during the period from 1942-44. (Retained - 31)
 Chi = 4.2, D = .57
55. YES ? NO There was at least one month when the population decreased during 1944-54. (Retained - 32)
 Chi = 1.52, D = .54
56. YES ? NO The 1934 population is smaller than the 1944 population. (Discarded)
 Chi = 0, D = .96

Directions: Study the following figure carefully. In front of the statement below it ENCIRCLE:
 YES, if the statement is supported by the data in the figure,
 NO, if the statement contradicts the data in the figure,
 ?, if the figure gives no evidence of the statement's being true or untrue.

(Adapted from Mussen and Conger, Child Development and Personality)

TABLE OF
 AVERAGE GAIN IN HEIGHT AND WEIGHT OF BOYS AND GIRLS FROM BIRTH
 TO FOUR YEARS



57. YES ? NO 2-year old girls weigh, on the average, 6 pounds.
 (Retained - 35)
 Chi = 1.25, D = .28
58. YES ? NO Physical development in the second year proceeds more slowly than in the first year.
 (Discarded)
 Chi = .546, D = .86
59. YES ? NO Between 1 and 2 years of age, the child's motor development is faster than his language development.
 (Retained - 33)
 Chi = 1.64, D = .86
60. YES ? NO At the end of 4 years the average girl is approximately 20 inches taller than when she was born.
 (Retained - 34)
 Chi = 1.19, D = .36

SECTION J

Read the following passage carefully:

A technician was smoking while preparing a gas in his laboratory. Suddenly there was a violent explosion.

Directions: Depending only on the information given above, ENCIRCLE
T, if the statement is true,
F, if it is false,
?, if it is uncertain.

61. T ? F The gas was inflammable. (Retained - 40)
Gri = 3.04, D = .54
62. T ? F The gas was not inflammable. (Retained - 58)
Gri = 2.76, D = .64
63. T ? F The technician was not smoking near the materials which exploded
(Retained - 39)
Gri = 2.76, D = .64
64. T ? F The gas made an explosive mixture with air (Retained - 37)
Gri = 2.96, D = .72
65. T ? F The explosion was due to the ignition of dynamite.
(Retained - 36)
Gri = 1.75, D = .75

SECTION K

Directions: Organize the following set of data so as to fit the skeleton outline below. In each of the blanks, write the number of the statement that best fits the pattern of the blanks: that is, in the first blank, write the number of the statement which you think should be the title, and so on.

TITLE: " _____ " .

I _____

A. _____

B. _____

II _____

A. _____

B. _____

III _____

A. _____

B. _____

- | | |
|---|---------------------|
| 66. Skeletal Growth. | (Retained - 45) |
| | Chi = 2.96, D = .72 |
| 67. Intelligence and the I.Q. | (Retained - 43) |
| | Chi = 1.97, D = .82 |
| 68. Emotional States and their Identification. | (Retained - 41) |
| | Chi = 1.77, D = .86 |
| 69. Learning vs Inheritance of Fears. | (Retained - 46) |
| | Chi = 2.96, D = .72 |
| 70. Mental Growth and Development. | (Retained - 47) |
| | Chi = 2.96, D = .72 |
| 71. Emotional Growth and Development. | (Retained - 49) |
| | Chi = 3.20, D = .68 |
| 72. Physical, Mental and Emotional Development. | (Retained - 42) |
| | Chi = 1.77, D = .86 |
| 73. Changes in Glandular Secretions. | (Retained - 50) |
| | Chi = 2.42, D = .68 |
| 74. Physical Growth and Development. | (Retained - 48) |
| | Chi = 2.96, D = .72 |
| 75. Developing Problem-Solving Abilities. | (Retained - 44) |
| | Chi = 2.52, D = .79 |

APPENDIX B

The Revised Test

FIRST REVISION OF
EXPERIMENTAL
GRADUATE APTITUDE TEST

1960

Department of Education, American University of Beirut.

SECTION A

Directions: ENCIRCLE THE LETTER OF THE WORD OR PHRASE THAT BEST COMPLETES EACH OF THE FOLLOWING:

Example

To be tidy means to be

- A transparent.
- B neat.
- C dirty.
- D rough.

-
1. The embryo in the case of human beings is formed from
 - A the egg or the ovum of the mother
 - B the sperm of the father
 - C both the egg and the sperm.
 2. To be tentative means to be
 - A conclusive.
 - B provisional.
 - C authentic.
 - D appropriate.
 3. Suppose you are asked to locate a particular author's magazine article in your term paper. The best way to find his article will be to look in the
 - A Reader's Guide to Periodical Literature.
 - B Book Review Digest.
 - C Library's main card-index under the name of the article.
 4. Taxes are charged mainly to
 - A provide revenue for the government.
 - B check inflation.
 - C guard against accumulation of money in the hands of a few persons.
 5. To recapitulate means to
 - A summarize.
 - B seize.
 - C catch hold of.
 6. In weather forecasting the most important factor a meteorologist has to consider is change in the air's
 - A humidity.
 - B pressure.
 - C temperature.

7. A pair of identical twins is formed as a result of the

- A fertilization of one egg by two sperms.
- B fertilization of two eggs by one sperm.
- C division of a sperm or an egg before fertilization.
- D division of one egg after fertilization.

8. Your teacher advises you to quote a particular author's magazine article in your term paper. The best way to find his article will be/look in the to

- A Education Index.
- B Encyclopedia Britanica.
- C Library's main card-index under the name of the article.

SECTION B

Directions: ENCIRCLE THE WORD IN THE PARENTHESES THAT BEST COMPLETES EACH OF THE FOLLOWING SENTENCES:

Example:
He came to see (I)
(me)

9. It was (her)
(she) that went out just now, wasn't it?

10. (Its)
(It's) too bad that your dog has hurt (its)
(it's) leg.

SECTION C

In the example below, the first word - Eyes - in the row of three words, is related to 'see' in the row in a certain way. 'hear' (encircled) in the list of five words below the row is related to 'ears' - the third word in the row in the same way as 'Eyes' are related to 'see' in the row.

Directions: ENCIRCLE THAT WORD IN THE LIST OF FIVE WORDS, WHICH IS RELATED TO THE THIRD WORD IN THE ROW IN THE SAME WAY AS THE SECOND IS RELATED TO THE FIRST IN THE ROW in each of the following items:

Example:

Eyes : see :: ears :

person hand face smell hear

11. Square : area :: cube :

six sides box volume three sides mass

12. Virtue : vice :: famous :

glorious evil popular notorious nice

13. Foot : leg :: hand :

fingers touch arm nails wrist

SECTION E

Directions: INSERT PUNCTUATION MARKS AND ADD OR REMOVE CAPITAL LETTERS
WHERE NECESSARY IN EACH OF THE FOLLOWING SENTENCES:

15. He washed examined his somewhat too easily growing beard and decided
to shave again.

16. An exhibit will begin on monday afternoon at half past one everybody
is invited the exhibit is free.

SECTION F

Directions: ENCIRCLE T IF THE STATEMENT IS TRUE OR F IF IT IS FALSE:

Example:

F Grass is green.

17. T F I.C.A. is a U.N. agency.

18. T F Red China is another name for Nationalist China.

19. T F Chronology is the science dealing with the measurement of time.

SECTION G

Directions: In each of the following items, the third statement may be true or false, depending upon the two statements preceding it, which you must assume to be true.

ENCIRCLE T IF THE THIRD STATEMENT IS TRUE OR F IF IT IS FALSE:

Example:

All men are mortal
He is a man
(T) F Therefore he is mortal.

20. All fruits have skins
Nuts are fruits
T F Therefore all nuts have skins.

21. Some men live to eat
He is a man
T F Therefore he lives to eat.

22. All toys are made of wax
This object is made of wax
T F Therefore this object is a toy.

23. All plants contain chlorophyll
Mushrooms do not contain chlorophyll
T F Therefore mushrooms are not plants.

24. All plocs nepp smoothly
All smooth neppers are ponns
T F Therefore all plocs are ponns.

SECTION H

Directions: Read the following passage carefully. In front of the statements below it ENCIRCLE:
YES, if the statement is supported by the data in the passage,
NO, if the statement contradicts the data in the passage,
?, if the passage gives no evidence of the statement's being true or untrue.

(Adapted from TIME, Atlantic Edition, November 16, 1959, p. 55)

It will take a generation or more to clear the state mental hospitals of the backlog of patients permanently crippled by old-time procedures that, far from making them better helped to make them worse. But seclusion rooms are being converted into kitchenettes and beauty parlors; camisoles and straps are disappearing. Shock treatment is seldom used, and only for selected patients. Though admission rates are rising, release rates are rising faster, so that in many states there is a net decrease in the numbers of mentally ill confined to hospitals.

25. YES ? NO Shock therapy is now used for some patients.
26. YES ? NO Admission rates are rising faster than release rates.
27. YES ? NO Throughout the country the proportion of mentally ill persons outside the hospitals is increasing.
28. YES ? NO "make them worse" (line 3 of the passage) refers to the physical condition of patients.

Directions: Read the following passage carefully. In front of the statements below it **ENCIRCLE:**

YES, if the statement is supported by the data in the passage,

NO, if the statement contradicts the data in the passage,

?, if the passage gives no evidence of the statement's being true or untrue.

(Adapted from Jones and Darkenwald, Economic Geography.)

The rate of population growth has not been uniform in all parts of the world. In some old lands population has declined, whereas in others, for example China and India, numbers have continued to multiply, reducing the standard of living to the subsistence level or a little above it. The population of France has been practically static for some time, even though there has been a marked change in the economic activities of the people. In the United Kingdom the population is gradually becoming stationary. In the United States the rate of increase is declining.

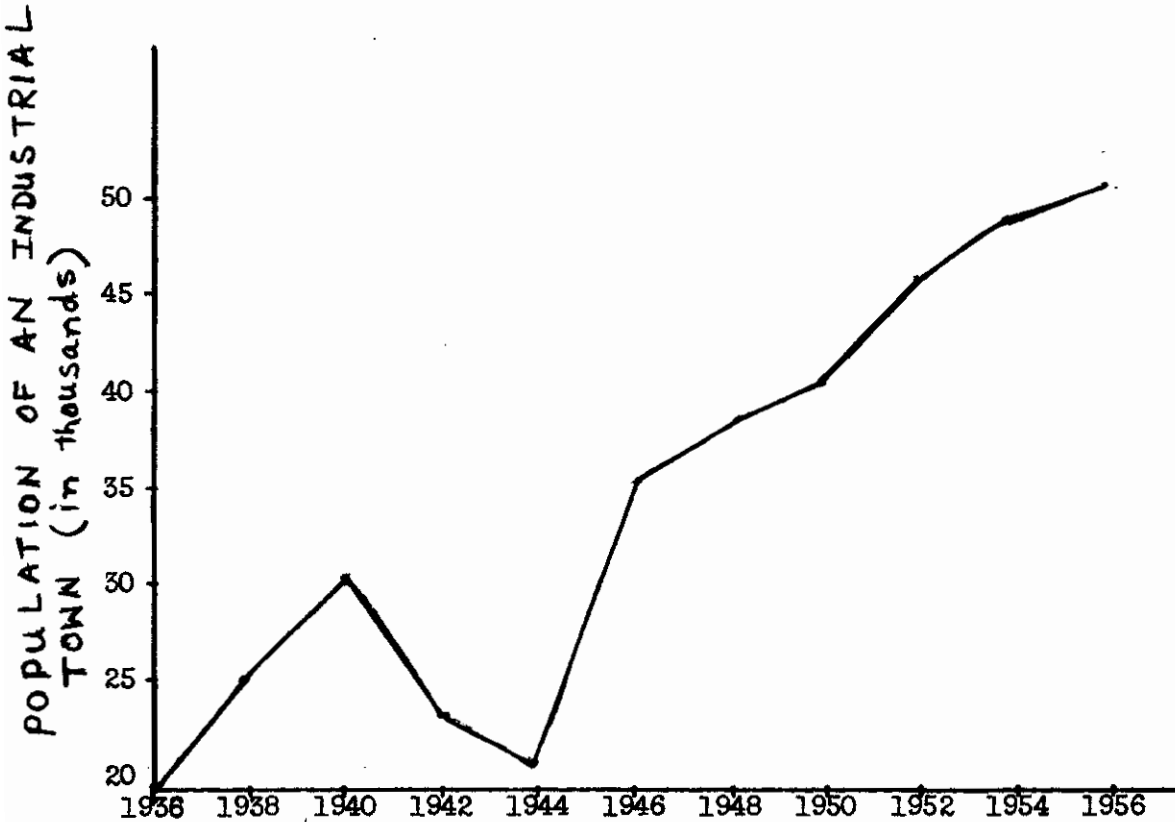
29. YES ? NO The population has increased by approximately equal amounts throughout the world.
30. YES ? NO In Pakistan, the standard of living is near the subsistence level due to overpopulation.

Directions: Study the following graph carefully. In front of the statements below it ENCIRCLE:

YES, if the statement is supported by the data in the graph,

NO, if the statement contradicts the data in the graph,

?, if the graph gives no evidence of the statement's being true or untrue.

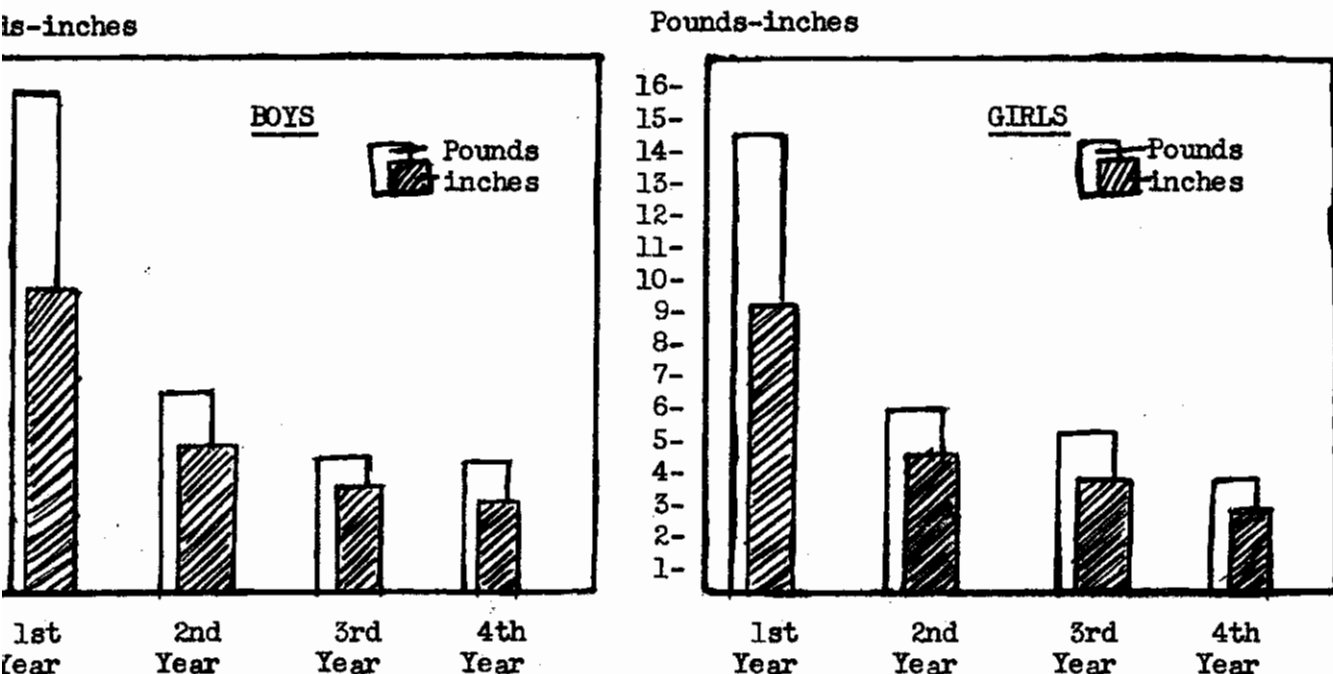


31. YES ? NO Many people died during the period from 1942 - 44.
32. YES ? NO There was at least one month when the population decreased during 1944 - 54.

Directions: Study the following figure carefully. In front of the statement below it ENCIRCLE: YES, if the statement is supported by the data in the figure, NO, if the statement contradicts the data in the figure, ?, if the figure gives no evidence of the statement's being true or untrue.

(Adapted from Mussen and Conger, Child Development and Personality).

TABLE OF
AVERAGE GAIN IN HEIGHT AND WEIGHT OF BOYS AND GIRLS FROM BIRTH TO FOUR YEARS



33. YES ? NO Between 1 and 2 years of age, the child's motor development is faster than his language development.
34. YES ? NO At the end of 4 years the average girl is approximately 20 inches taller than when she was born.
35. YES ? NO 2-year old girls weigh, on the average, 6 pounds

SECTION I

Read the following passage carefully:

A technician was smoking while preparing a gas in his laboratory. Suddenly there was a violent explosion.

Directions: Depending only on the information given above, ENCIRCLE
T, if the statement is true,
F, if it is false,
?, if it is uncertain.

36. T ? F The explosion was due to the ignition of dynamite.
37. T ? F The gas made an explosive mixture with air.
38. T ? F The gas was not inflammable.
39. T ? F The technician was not smoking near the materials which exploded.
40. T ? F The gas was inflammable.

SECTION J

Directions: Organize the following set of data so as to fit the skeleton outline below. In each of the blanks, write the number of the statement that best fits the pattern of the blanks: that is, in the first blank, write the number of the statement which you think should be the title, and so on.

TITLE: " _____ ".

I _____

A. _____

B. _____

II _____

A. _____

B. _____

III _____

A. _____

B. _____

41. Emotional States and their Identification.
42. Physical, Mental and Emotional Development.
43. Intelligence and the I.Q.
44. Developing Problem-Solving Abilities.
45. Skeletal Growth.
46. Learning vs Inheritance of Fears.
47. Mental Growth and Development.
48. Physical Growth and Development.
49. Emotional Growth and Development.
50. Changes in Glandular Secretions.

APPENDIX C

Steps Leading to the Degree of Master of Arts for
Students from other Colleges and Universities,
Entering the Department of Education, American
University of Beirut. (Department of Education,
American University of Beirut, July 1959).

Steps Leading to the Degree of Master of Arts
for Students from other Colleges and Universities Entering the
Department of Education, American University of Beirut

1. Candidate applies for admission to the University Administration through the Office of the Registrar. Registrar's evaluation of credentials and the result of the English Entrance Examination determine the level at which candidate may undertake work.
2. If the Registrar's evaluation determines the candidate's eligibility to work for the M.A. degree in education, the candidate's records are examined by the Department of Education. Official acceptance is requested from the Committee on Graduate Study for those candidates who are deemed promising.
3. A program of courses for the first semester is worked out in consultation with the Departmental advisor to graduate students. Not less than 12 and not more than 18 credit hours may be taken by full-time students in one regular semester. A minimum of 6 and a maximum of 9 credits may be taken by full-time students during the summer session.
4. At the end of the first semester of work the student sits for the Written and Oral Qualifying Examination of the Department. The purpose of this examination is to determine whether the student has the ability to undertake independent research and complete successfully the graduate work leading to the Master of Arts degree. This examination attempts to test the candidate's ability to
 - A. Express himself adequately in written and spoken English
 - B. Organize data in a logical, coherent manner
 - C. Perceive relevant aspects of a problem
 - D. Devise relevant research techniques for studying the problem
5. A departmental decision is then made regarding the student's continuation in the graduate program, and, if he is to continue, the number of credit hours and the minimum length of residence required for the award of the M.A. degree. This decision is based on the following factors:
 - A. The results of the Written and Oral Qualifying Examination
 - B. The student's record in course work during the first semester
 - C. The student's previous degrees, the quality of his previous academic work and his teaching experience.

6. An accepted graduate student, who has fulfilled all the undergraduate requirements for admission, is required to spend a minimum of one full academic year and one summer session, and to complete 30 credit hours, including 6 for the thesis. Usually, however, this minimum requirement is extended, depending on the judgment of the Department of Education. Credit hours required for graduation may thus range between 30 and 60 and the period of required residence may extend from two semesters and one summer to four semesters and one summer.
7. The student's course work must include a minimum of 10 credits of courses numbered 300 and above. For graduation his average in such courses, regardless of the number he takes, must be 80 or above. In addition, his average in all courses taken together must be 75 or above.
8. One semester before his earliest graduation date at the latest, a student must make a successful oral defense of a proposed thesis outline and prospectus. An examining committee and thesis chairman have been appointed by this time. At about this date he sits for the Graduate Written Comprehensive Examination which is designed to test the candidate's broad comprehension of the field of education and to examine his ability to apply his learnings in the various fields to the task of producing a unified thesis. If his oral defense of outline and prospectus is acceptable, and if he passes the Graduate Written Comprehensive Examination, the student is recommended to candidacy for the degree of Master of Arts.
9. Not later than three weeks before the expected graduation date, the student must present one original and two carbon copies of his thesis to the office of the Dean of the School of Arts and Sciences. These copies then become available for purposes of the oral examination.

F. R. Korf
Acting Chairman
Department of Education

July, 1959

WORKS CITED

Books

- Davis, Frederick B. "Item Selection Techniques," Educational Measurement, ed. E.F. Lindquist. Washington, D.C.: (American Council on Education), 1951, pp. 266 - 328.
- Freeman, Frank S. Theory and Practice of Psychological Testing. New York: (Henry Holt and Company), 1950.
- Guilford, J.P. Fundamental Statistics in Psychology and Education. New York: (McGraw-Hill Book Company), 1950.
- Mosier, Charles I. "Batteries and Profiles," Educational Measurement, ed. E.F. Lindquist. Washington, D.C.: (American Council on Education), 1951, pp. 764 - 808.
- Skinner, C.E. Essentials of Educational Psychology. Englewood Cliffs, N.J.: (Prentice Hall), 1958.
- Thorndike, Robert L. "Reliability," Educational Measurement, ed. E.F. Lindquist. Washington, D.C.: (American Council on Education), 1951, pp. 560 - 520.
- Thorndike, Robert L, and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education. New York: (John Wiley and Sons), 1955.
- Travers, R.M.W. Educational Measurement. New York: (Macmillan Company), 1955.

Unpublished Material

Ahmad, Nasim. Problems of Pakistani I.C.A. Students at A.U.B.
(unpublished M.A. Thesis, American University of Beirut),
1959.

American University of Beirut. Department of Education. Steps Leading to the Degree of Master of Arts for Students from other Colleges and Universities Entering the Department of Education, American University of Beirut. (Mimeographed announcement), 1959.

SOURCES CONSULTED

Thorndike, Robert L. Personnel Selection. New York: (John Wiley & Sons), 1949.

Travers, Robert M.W. How to Make Achievement Tests. New York: (The Odyssey Press), 1950.