T
10

ADAPTATION OF HENMON-NELSON TEST OF MENTAL ABILITY TO SYRIAN STUDENTS IN THE SECONDARY SCHOOL

by

BALKIS ABDUL KADER AWAD

A Thesis
Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Arts in the
Departments of Psychology & Education
American University of Beirut.

September 1951

Dedicated to my Parents

I wish to express my deep gratitude and thanks to the Chairman and members of the Departments of Education and Psychology at the American University of Beirut for their help and guidance in making this work possible.

TABLE OF CONTENTS

ADAPTATION OF HENMON-NELSON TEST OF MENTAL ABILITY TO SYRIAN STUDENTS IN THE SECONDARY SCHOOL.

The success of an educational system is measured by the degree to which it has been able to discover the abilities and needs of pupils in order to provide for them the adequate type of education. Therefore many investigations in the West were made and different types of tests of a sound scientific basis emerged.

Examinations in Syria are still of the traditional type and do not serve the real aim of education. Moreover, the abolition of fees in the Syrian Secondary Schools intensified the need for new and better types of examination. Consequently the purpose of this study is to prepare one type of the testing tools applied in the West - The Intelligence Test. The shortest way for preparing this test, in the writer's opinion, is to adapt a good one used in the West; Henmon-Nelson Test of Mental Ability was chosen for this purpose.

Hence the text and the manual of the test were translated and adapted to Arabic language and Syrian environment.

Modifications introduced in the content of the test are due to two general causes:

a) The inadequacy of facts and information to the new environment.

b) The different syntax in the two languages.

Alterations introduced in the manual are in three aspects:

a) in the number of exercises,

b) in the content of directions,

c) in the time limit of the test.

The sample of this study is random and composed of one-hundred and forty-two students (72 girls and 70 boys) of Arts and Science divisions in the fifth year, in two public secondary schools in Damascus.

The score of every student in this test is the total of the right answers and the student's status is judged by ranking him only approximately according to his group, because all the norms tables found in the original manual are based upon representative students living in U.S.A. Thus they cannot be applied to Syrian students.

Statistical techniques, employed in studying the results of the adapted test, revealed that twenty items were ineffective and must be eliminated from the test.

The reliability of the test, even after the elimination of the inadequate items, was high and equal to .911.

The validity of the test was computed by correlating the test scores with the students' marks in the Mid Year School Examination on the one hand, and with the Brevet Averages on the other hand, treating every one of the eight divisions apart. Few subject matters were selected for the averages representing the teachers' marks in the Mid Year School Examination so as not to go in the long marks weighting process. The coefficients of correlation were ranged from .04 to .54. Three out of eight were very low (.04, .08, .12). Since no dependable valid criteria

were available to compare the test with and since Teachers' marks and Brevet Averages cannot be relied on very much due to their deficiency in many respects, therefore we cannot fairly judge the low value of the actual coefficients of validity.

The Arithmetic Mean of the group in this study is 50.70, S.E. = 1.12; the Median = 49.5, S.E. = 1.41 and the Standard Deviation is 13.5, S.E. = .79.

Shortly after reviewing this study one comes to the conclusion that this Adapted Test in its actual content although representing a step in the right direction, is in need of further study and refinement to become a standardized test.

CHAPTER I

INTRODUCTION

THE SYSTEM OF EXAMINATIONS IN SYRIA

Examinations in Syria as in several other countries in the world, are a problem which must be reviewed and studied deeply in the light of the new concepts of education.

They are rigidly prepared and administered and they are stressed in every period of study. No student can pass from one class to the next higher without first passing the examinations in the previous stage. Side by side with the examinations held usually by teachers in the school in the different stages of the educational ladder there are three main public examinations which are prepared under strict secrecy by the Committee on Education and Instruction of the Ministry of Education. At the end of the fifth year in the elementary cycle there is "the Public Primary Examination" or the "Certificate Examination". The secondary education consists of 6 years of study. It is divided into two stages; the first four years compose the"intermediate" stage and terminate in the Public Intermediate Examination, "The Brevet"; the last two years, the "preparatory" stage, are divided into literary and scientific sections, with different curricula. This stage leads to a public examination after which the successful pupil gets the Certificate of Completion of Preparatory or Secondary Studies or "The Baccalauréat". Candidates

for each of these examinations must be not less than a certain definite age[1] and no pupil can sit for the final secondary examination until two years after he has taken the Certificate of the intermediate stage (The Brevet). All these examinations are written and of essay type. One primary school examination session is held each year but two chances are provided each year for the students of the intermediate and preparatory stages, one in June and one in September under certain conditions.

Viewed from the stand point of students and parents the intermediate school examination and the secondary school examination are regarded as serious obstacles to be overcome because their passing is an essential prerequisite, not only for admission to higher institutions but to many intermediate careers. Consequently pupils are subjected not only to the pressure of the schools themselves, but also to family, social and economic pressure which tends to reinforce the intellectualistic atmosphere of the school. Moreover the quality of a school has been judged by the number of its successful pupils. Therefore there is a great fear of the examinations not only on the part of students but by teachers also. Consequently many schools and parents try to provide special classes outside regular school

---

(1) For example no student can sit for the Certificate Examination if he is less than 10 years old and for the Brevet if he is less than fourteen and for the Baccalaureat if he is under sixteen. These conditions are only for students who follow their study regularly in schools. Those who are not registered in private or public schools must be respectively not less than 13, 17 and 19. Anzimat Al-Imtihanat Al-'Ammah, The Department of Education, Syrian Republic, Damascus, 1945, Decrees, 149, 150, 151.

انظمة الامتحانات العامة ـ وزارة المعارف ـ الجمهورية السورية ـ دمشق سنة ١٩٤٥ ـ مرسوم رقم ١٤٩ ـ ١٥٠ ـ ١٥١ .

hours in order to prepare students for these examinations. On the other hand students on the whole do not pay much attention to courses that are not included in the public examinations, but they spend their time in cramming and memorizing the facts found in the text books designed to assist in the examination. Therefore, school education becomes to a large extent a preparation for the examinations, and the examinations come to be an end in themselves and act as a hindrance to educational experimentation and progress. We add to these characteristics another outstanding feature of these examinations and this is the great degree of unreliability in the assessing of results so that the fate of the candidates in the examination depends largely upon chance.

We can say briefly that the examinations in Syria are injurious to the health of the best students and lead to overstrain and worry and often to despair and a feeling of inferiority in hundreds of pupils. Moreover they encourage faulty methods of study and superficial learning, things which cannot go with modern aims in education.

Examinations in Syria serve a dual purpose; they are used both as tests of achievement and as methods of selection which becomes more intense as the time of admission to the universities approaches. Therefore all those whose abilities and aptitudes do not fit them into the traditional mould are excluded.

The problem of examinations is a common problem in the world of education despite the diversity of the political and

economic systems. However it cannot be studied apart from society, nor can it be understood if approached only from the point of view of techniques of evaluating achievement or of setting educational standards. The use of examinations is determined not only by a philosophy of education but also by a social philosophy and the demands of the present civilization. Therefore it cannot be attacked from a unitary point of view and each country must develop its own resources and methods for investigating the problem. The technical, psychological and statistical studies of examinations are as much concerned with the needs of society as of the individual and these studies will remain meaningless unless they are guided by a philosophy of educational values and purposes.

Consequently many studies in our country must be centered around the problem of overcrowding of students in the universities and the governmental professions, the educational and vocational maladjustments, taking into consideration the urgent needs of the Syrian society in this stage of civilization. Our changing world demands on the one hand an increase and extension of educational opportunities and on the other hand the discovery of the right education for the right individual, and beyond this the urgency of devising methods which may reduce the intensity of the current pressure to academic and governmental professions, and to encourage technical ability and manual labor which are still too often regarded as of an inferior value. This is what is badly needed in Syria at present especially since the

abolition of fees in the secondary schools has intensified the problem of selection and distribution. A system of free education implies selection and distribution of pupils and orientation or guidance. Therefore the word "failure" will have no place in a comprehensive philosophy of education. Such eliminated individuals deserve something more constructive and helpful than merely a failing grade. What is obvious at this point is that if secondary education is to be provided for all, then it cannot be the same for all. The fear has been expressed that to adapt the curriculum to the students' abilities and aptitudes means the lowering of standards, while, in fact the contrary is true. When all the different students are submitted to the same curriculum the standard becomes lower because most of the students in this case are not carried along the lines on which they are best fitted to advance. Education in Syria must meet the highly differentiated interests and abilities of the large body of adolescents for which new educational opportunities must be planned. We quote Professor C. Delisle Burns at the conference on Education held at Eastbourne in 1931, "... you need a new kind of examination system, one based rather on the social assumption that you want distributions of functions within each section of society. It is much less important to see who is at the top and who is at the bottom. There is no top and bottom really; society is simply an interlocking of certain structures almost organic structures - in regard to which you cannot say that one part of the structure is purely directive and the other

part is purely instrumental.[1]

One may conclude from the study of Syrian examinations that their purpose is mainly to discriminate between students who should pass or fail on the basis of some preconceived standard of achievement in order to standardize the instruction of the schools or other educational institutions. While the chief purpose of examinations must be to diagnose pupils' difficulties and to improve methods of instruction and guidance. The success of an educational system must be measured by the degree to which it has been able to discover the abilities and needs of pupils and to provide for them the type of education from which they can profit at their best. We quote Sir Graham Balfour on this point, "to enable the right pupils to receive the right education from the right teachers at a cost within the means of the State under conditions which will enable the pupils best to profit by their training."[2]

To discover the abilities and needs of pupils many investigations in the West were made to construct more scientifically accurate, reliable, and valid types of examinations than the traditional type. The chief characteristic of the new type test is that it calls for a short answer by one or a few words, by a check mark, or by a number. In place of the small number of questions which may be asked in the essay type, 50 to

---

(1) L. Kandel, Examinations and their substitutes in the United States, The Carnegie Foundation for the advancement of teaching, New York 1936, P. 72.

(2) Ibid., P. 162.

100 or more items may be answered in an hour's new type examination. From this point of view alone the new type test reduces the element of chance or luck, so far as the examinee is concerned.

Generally the new type test is considered superior to the old for a variety of reasons, among them the following:

1. It measures to a known degree what it is designed to measure.
2. It is statistically more reliable.
3. It is more objective since subjective factors such as bias, prejudice, personal opinions, and temperament are reduced to a minimum and the scores would always be the same whatever the number of examiners might be.
4. It is more comprehensive because it includes a wider sampling of the subject of the examination.
5. It is easily administered and marked with economy of time and effort for both examiners and pupils.
6. The results can be interpreted more easily and used for specific purposes of diagnosing weaknesses or encouraging ability.

However it is admitted by those who can speak with authority that these tests themselves are not yet perfect. Therefore a better examination is a combination of the new and old type of questions. At the same time efforts must be encouraged to make the scoring of the essay type as reliable and valid as possible and to devise and use objective scholastic, aptitude tests and intelligence tests to supplement the regular examina-

tion with the attention always directed to what is missed rather than how much.

In other words no test, however perfect can tell everything that is needed about a pupil. Therefore the collection of as much information as possible about the intellectual, personal, and social characteristics of an individual both in and out of school are necessary for his guidance. In this collecting process both the individual and the society of which he is a member are considered.

Thus information about any pupil must be comprehensive and continuous and this is provided in the cumulative record which is one of the most important outcomes of the testing movement. It becomes now indispensable to facilitate educational guidance. For example the Secondary School Cumulative Record Form provides for the recording of student's information on the following items: Name, religion, sex, date of birth; mental age, chronological age, intelligence quotients, school grade achieved; school grade attended; achievement test and school marks; height and weight, photograph; schools attended; record of attendances and absences; causes of absence; discipline; unusual accomplishments; mental, emotional, and physical experiences; extra-curricular experiences, athletic and non-athletic; clubs and offices; vocational and professional preferences; interests reported; special defects; health; mental hygiene; social adjustments and home conditions; personality ratings and measurements.

Information on these items is recorded and graphically

represented so that the pupil's progress can be traced easily across the record form as it is shown in photograph I and II.(1)

Hence the different tests according to the new testing program are used for the pupil rather than on him. In such a program, elements of marking, competition and imposition from above are minimized and in the most successful programs the pupil takes the tests because he wants to know the results. Also the new testing program brings with it an increasing criticism against the prevailing "time" concept in the school according to which the student is obliged, as in Syria, to spend certain definite years in school though he is capable of passing in a shorter time than the normal student.

Going back to Syria, we believe that to improve the system of examinations apart from the other interrelated problems such as a definite philosophy of education based on a deeper study of the requirement of society and individual together, cannot be achieved. Therefore we see that the whole educational system must be reviewed and redefined, and consequently the testing instruments and measurement program come naturally under this survey.

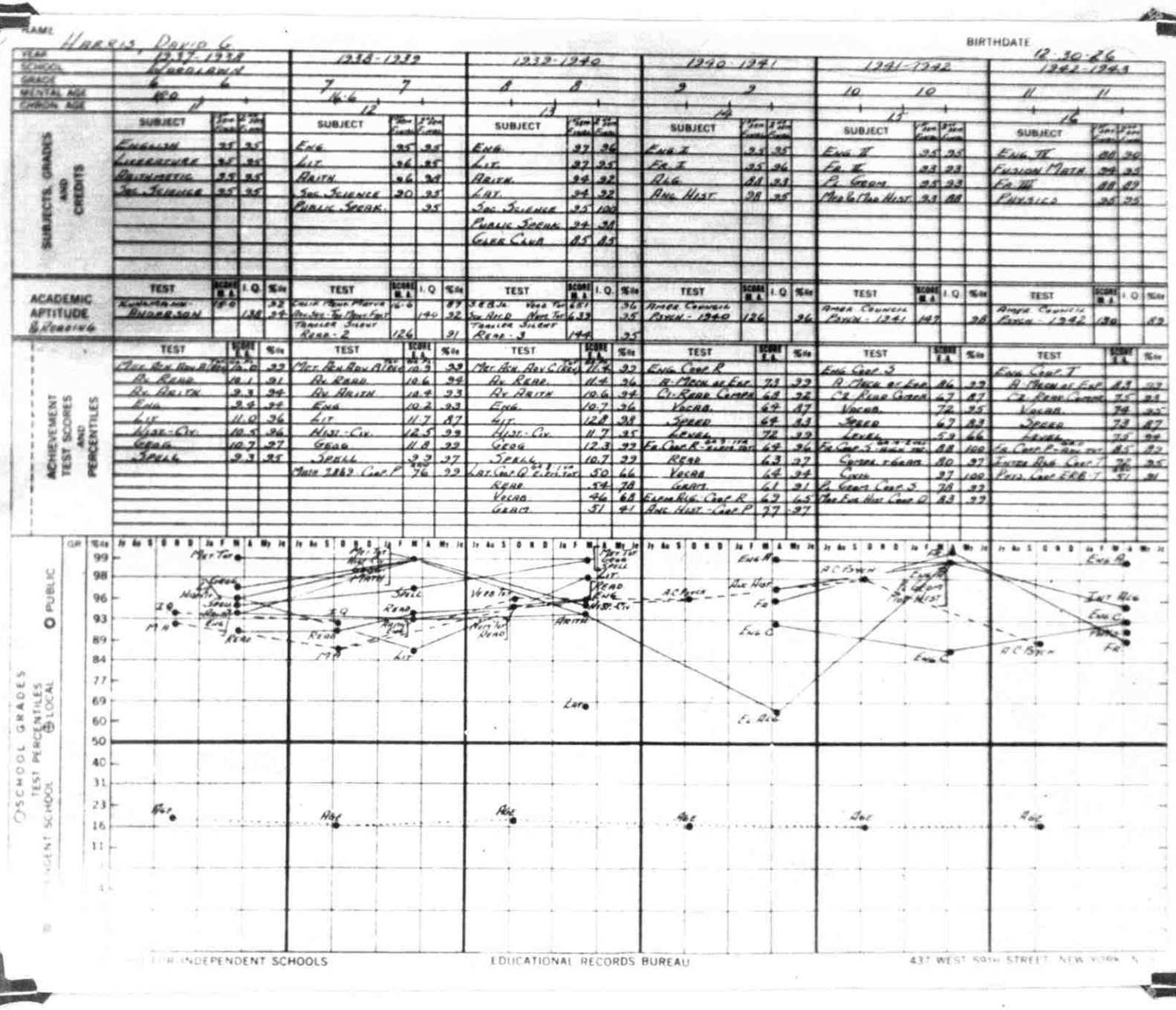
However to offer to the Ministry of Education some new type of objective tests which are constructed on a sound scientific basis may stimulate the efforts and direct the attention of teachers and those who are interested in education toward this important point.

---

(1) Arthur E. Traxler, "Techniques of Guidance", Harper and Brothers, New York, 1945, P. 216.

Our present study, as will be shown in the following chapters, is an attempt to prepare one of the most important instruments used in measuring students' mental ability - the intelligence test.

Photograph I - A Cumulative Record Form Used in the United States for One Pupil (front).



NAME Harris, David G — BIRTHDATE 9-30-26

	1937-1938	1938-1939	1939-1940	1940-1941	1941-1942	1942-1943
YEAR	1937-1938	1938-1939	1939-1940	1940-1941	1941-1942	1942-1943
SCHOOL	Woodland					
GRADE	6 6	7 7	8 8	9 9	10 10	11 11
MENTAL AGE	12-0 +	14-6 +	13 +	14 +	15 +	16 +
CHRON. AGE	12	12				

SUBJECTS, GRADES AND CREDITS

SUBJECT	Term Grade	Term Cum.
1937-1938		
English	95	95
Literature	95	95
Arithmetic	95	95
Soc. Science	95	95
1938-1939		
Eng.	95	95
Lit.	96	95
Arith.	96	95
Soc. Science	90	95
Public Speak.		95
1939-1940		
Eng.	97	96
Lit.	97	95
Arith.	94	92
Lat.	94	92
Soc. Science	95	100
Public Speak.	94	98
Glee Club	85	85
1940-1941		
Eng. I	95	95
Fr. I	95	96
Alg.	88	93
Anc. Hist.	98	95
1941-1942		
Eng. II	95	95
Fr. II	93	93
Pl. Geom.	95	93
Mod. & Ther. Hist.	93	88
1942-1943		
Eng. III	90	90
Fusion Math	99	95
Fr. III	88	89
Physics	95	95

ACADEMIC APTITUDE & READING

TEST	Score M.A.	I.Q.	%ile
1937-1938			
Kuhlmann-Anderson	150	138	99
1938-1939			
Otis Higher Mental Exam.	146	140	99
Thurstone Test of Mental Factors			97
Thurstone Silent Read. - 2	126		91
1939-1940			
S.R.A. Jr. High Test	187		96
Sun Dial Mental Test	6.33		95
Thurstone Silent Read. - 3	144		95
1940-1941			
Amer. Council Psych. - 1940	126		96
1941-1942			
Amer. Council Psych. - 1941	147		98
1942-1943			
Amer. Council Psych. - 1942	130		83

ACHIEVEMENT TEST SCORES AND PERCENTILES

TEST	Score M.A.	%ile
1937-1938		
Met. Rea. Rev. Abbrev.	10.0	99
Rv. Read.	10.1	91
Rv. Arith.	9.3	94
Eng.	9.4	94
Lit.	11.0	96
Hist.-Civ.	10.5	96
Geog.	10.7	97
Spell.	9.3	95
1938-1939		
Met. Rea. Rev. Abbrev.	10.9	99
Rv. Read.	10.6	94
Rv. Arith.	10.4	93
Eng.	10.2	93
Lit.	11.7	97
Hist.-Civ.	12.5	99
Geog.	11.8	99
Spell.	9.3	97
Metr. 1942 Conc. P	76	99
1939-1940		
Met. Rea. Rev. Abbrev.	11.4	99
Rv. Read.	11.4	96
Rv. Arith.	10.6	94
Eng.	10.7	96
Lit.	12.8	98
Hist.-Civ.	11.7	95
Geog.	12.3	99
Spell.	10.7	99
Lat. Cae. Q. Elem. Test	50	66
Read.	54	78
Vocab.	46	68
Geog.	51	41
1940-1941		
Eng. Conc. R		
R. Mech. of Eng.	73	99
Ct. Read. Comp.	68	98
Vocab.	69	97
Speed	69	83
Level	72	99
Fr. Conc. R. Elem. Test	69	96
Read.	63	97
Vocab.	64	94
Geog.	61	91
Elem. Alg. Conc. R	69	65
Anc. Hist. - Conc. P	77	97
1941-1942		
Eng. Conc. S		
R. Mech. of Eng.	86	99
Ct. Read. Comp.	67	87
Vocab.	72	95
Speed	67	83
Level	59	66
Fr. Conc. S. Elem. Test	88	100
Comp. & Gram.	80	97
Conc.	97	100
Pl. Geom. Conc. S	78	99
Ther. Eve. Hist. Conc. O	83	99
1942-1943		
Eng. Conc. T		
R. Mech. of Eng.	83	99
Ct. Read. Comp.	75	94
Vocab.	74	95
Speed	73	87
Level	75	94
Fr. Conc. P. Elem. Test	85	99
Comp. & Gram. Conc. T	76	95
Phys. Conc. ERB T	51	91

SCHOOL GRADES ○ PUBLIC
TEST PERCENTILES ⊕ LOCAL
INDEPENDENT SCHOOL

EDUCATIONAL RECORDS BUREAU — 437 WEST 59TH STREET, NEW YORK, N.

Photograph II - A Cumulative Record Form Used in the United States for one Pupil (back).

	M F	BIRTHDATE—PLACE	GEN'L HEALTH	RELIGION	RACE OR NAT'L'TY	OCCUPATION	ADDRESS
CHILD Harris, David C.		4 Nov 26 Chicago	Good	Epis.	American	Student	1946 Maryland Avenue, Chicago, Ill.
FATHER Harris, Edward		7 18 97 Brooklyn NY	Excellent	"	"	Lawyer	RES. 1946 Maryland Avenue, Chicago, Ill. BUS. 724 S. Michigan Avenue, Chicago, Ill.
MOTHER Harris, Mary Jennings		2 11 98 Gary, Ind	Good	"	"	Housewife	RES. 1946 Maryland Avenue, Chicago, Ill. BUS.
STEP PARENT OR GUARDIAN							RES. BUS.

LANGUAGE SPOKEN IN HOME BEFORE 10— English AFTER 10— English

TYPE OF COMMUNITY BEFORE 10— Residential AFTER 10— Residential

IF PARENTS SEPARATED GIVE DATE

YEAR AND AGE	1937-38 11	1938-39 12	1939-40 13	1940-41 14	1941-42 15	1942-43 16
ADVISER	Saithers	Saithers	Jeffries	Jeffries	Haynes	Haynes
ATTENDANCE	A Regular T	A Regular T	A Regular T	A Regular T	A Regular T	A Regular T
DISCIPLINE	Excellent	Good	Good	Excellent	Good	Good
HOME INFLUENCES AND COOPERATION	Splendid	Good	Very Cooperative Parents	Good	Fine Home Background	Good
MENTAL AND EMOTIONAL	An excitable boy. Hampered in speech at first. Now intelligent and mature. A leader to the core.	Outstanding Student	Rather nervous and high strung, but has himself well under control.	Splendid Ability	Intelligent and well balanced.	Very high intelligence.
PHYSICAL AND ATHLETIC	Good Health. Interested in Athletics	Slight glandular trouble. Nerro glasses.	(Good) Eyes a little weak but ok on normal with glasses.	(Good) Health	Very good. Has developed a great deal physically this year.	Good. Takes an active part in Track, Baseball, Rifle Club.
EXTRA-CURRICULAR ACTIVITIES AND INTERESTS	Stamp collecting, Model ships, Planes, Books, Construction.	Made models of Battles of the U.S. Navy. Reads history of the Navy.	Model ships. Reading of military and naval history.	Model Aeroplanes, Reading.	Boat Building. Summer camp.	Boat Building, Shooting, Mountain Climbing.
NOTABLE ACCOMPLISHMENTS AND EXPERIENCES	Excellent Student. Advanced for age. The class leader by virtue of personality.	Leads the class in general ability.	A class leader.	Class Vice President.	Class President.	Elected to Student Council. Rifle Club Team.
EDUCATIONAL PLANS	St Paul's School	Naval Academy	Naval Academy.	College.	College.	Hopes to get into Navy College. Reasons after graduation or service here.
PERSONALITY RATINGS +2	Ability to lead, cooperate. Unselfish.	Splendid personal qualities.	Natural Leader.	Cooperation, Leadership.	Ability to get things done.	Fine personality.
+1						
0						
-1						
-2						

REMARKS An outstanding boy in every way. Excellent prospect for a good college. Educational plans very [illegible]. Some aspect of diplomatic service. Interested in [illegible]. Good material.

CHAPTER II

MENTAL TESTS USED IN THE WEST FOR THE EVALUATION OF THE INDIVIDUAL'S MENTAL ABILITIES

I - Kinds and Characteristics of These Tests:

The testing movement in the West, especially in America, touches on vital questions in all phases of life. Different agencies, educational, industrial, clinical and civic - regularly apply standard measures of behavior. Several millions of tests are used every year in the U.S.A. alone; which have a great economic significance in the selection of employees for various types of work and in the selection of students for various educational programs.

The purpose of every Mental test is to find out differences between individuals. This implies two functions: prediction and diagnosis.

Prediction refers to the differences between individuals or between an individual and some standard, while diagnosis, emphasizes differences among the various characteristics of the same individual.

Many various types of tests are usually employed in education and they may be classified on different bases according to their scope, content, administration etc. The following appears to be a satisfactory classification of the measurement instrument used in the ordinary school in U.S.A.

A. Oral

B. Written

1. Informal (non-standardized)

a. Traditional (essay type)

b. Objective (new type)

2. Formal (standardized)

a. Achievement

(1) General (survey)

(2) Specific (diagnostic, practice, etc.)

b. Intelligence

(1) General (individual and group)

(2) Specific (aptitude or prognosis)

c. Character and Personality. (1)

A formal or standardized test begins usually as an informal test but after it is subjected to a long experimental study and revisions it becomes standardized or formal.

The three types of standardized tests employed in testing students' mental abilities and achievement are - intelligence tests, aptitude tests and achievement tests. Though they are built on the same bases and principles yet each one of them has its own characteristics.

a) Aptitude Tests:

An aptitude test is a test which is designed to predict future success of the individual in some field. We quote Freeman,

---

(1) C.C. Ross, Measurement in Today's Schools, Printice-Hall, Inc. New York, 1941, P. 25.

"An aptitude is a condition or set of characteristics indicative of an individual's ability to acquire with training some specific knowledge, skill, or set of responses. An aptitude test, therefore, is a device designed to indicate a person's potential ability for performance of a certain type of activity of a specialized kind and within a restricted range."(1)

When we say that "X" has an aptitude for a certain type of activity such as musical aptitude, or mechanical aptitude we mean his capacity to acquire skill or proficiency in this activity if suitable conditions were provided.

Since the student himself is the important thing to the tester - aptitude test results must be used only as one kind of information in a total picture. Therefore it is essential that their results must be supplemented with school records, intelligence test, interviews and personality test, etc.

b) Achievement Tests:

A test of educational achievement is designed to measure knowledge, understanding, or skills in one subject such as history, mathematics or several subjects given in school. The number of these tests is very large. They cover almost every subject taught in elementary, secondary school and colleges in America. Some of them are devised for the purpose of testing the amount learned of the subject material, others have a diagnostic purpose.

Here is a summary of the most important purposes for

---

(1) F.S. Freeman, Theory and Practice of Psychological Testing, Henry Holt and Company, New York, 1950, P. 263.

which these tests are used:

1. They are often used to find out the effectiveness of different teaching methods.
2. They help teachers to evaluate their effectiveness when they are used side by side with tests of intelligence.
3. They are employed sometimes for the classification of students in relatively homogenous groups.
4. They help teachers and counsellors to diagnose each pupil's strength and weaknesses. For this reason, more and more schools are using achievement tests at the beginning of the school year, since there is little merit in testing after it becomes too late to profit from the results.
5. To select students by the institutions for particular types of education.
6. Their results are sometimes used by the clinical psychologist for students whose adjustment problems are connected with their deficiency in certain subjects.
7. They are useful in maintaining the standards of a school or for the evaluation of a school system from year to year.

Some of the criticism aroused against these tests is their emphasis on facts and information. "Facts and skills loom so large in the usual classroom that teachers and test designers have often emphasized them out of proportion to other types of outcome."(1)

---

(1) Lee J. Cronbach, "Essentials of Psychological Testing", Harper and Brothers, New York, 1949, P. 276.

The same meaning has been mentioned by Freeman, "Most tests of achievement are devoted very largely to the measurement of the amount of information recalled, or skills and techniques acquired. In more recent years, however, an increasing number have been devised to measure such educational results as problem solving, drawing inferences from subject matter, applying generalizations to specific situations and problems, attitudes and appreciations developed by the study of course materials, and practices and skills developed in the study of a given subject."(1)

Briefly in a sound measurement program, evaluation of a pupil or a teacher is never based upon the results of these tests alone; but the data of these tests must be added to other different sources in order to make a final evaluation of the individual.

II - <u>The Steps Followed in Constructing a Standardized Test.</u>

Constructing a standardized test is not easy at all. It demands from the person high practical and theoretical qualifications in this field. An understanding of the theoretical principles and assumptions upon which tests are constructed as well as preparation in statistics especially in statistical reasoning and methods, are essential for a test worker.

a. The first step in the development of a psychological test is to define what is to be measured, whether it is intelligence,

---

(1) Freeman, op. cit. P. 309.

specific aptitude or school achievement. Because any given test measures a limited aspect of the person being tested.

b. After a definition has been formulated, the psychologist begins to break down his definition into component processes. Binet, for example, gave the following as the components of the intelligence: memory, mental images, imagination, attention, comprehension, suggestibility, esthetic appreciation, muscular strength, etc... Spearman, Thorndike and Thurstone later found out more specific processes which according to them should be tested.

Then after analyzing his definition, the test builder, begins to select items which may represent well the traits or functions he has devised to sample. It is not necessary to present the testee with unlimited situations in which the traits are manifested. Therefore it is sufficient to select an adequate sampling representing the whole situations. Two steps are here followed for this purpose:

1) The test designer chooses the most prominent components of what is to be measured. In the case of intelligence for example he selects the constituents: vocabulary, arithmetical problems, verbal comprehension, analogies, etc.
2) He studies the operational levels within each constituent which must be selected, e.g. the kind of arithmetical processes, the kind of words involved in the vocabulary constituent, etc. ..

c. After the test builder puts his original items, he

subjects them to a long series of try-outs on the groups of the individuals for whom the test is devised. The results thus obtained are exposed to statistical analysis within the framework of a series of criteria of validity. As a result of this process some of the items are retained, others are rejected and some new ones are added.

d. These new added items with the previously selected ones in step (c) are again subjected to the same methods of validation many times before the final test is emerged.

This entire process of try-outs and the rigid statistical analysis and scrutiny within the criteria agreed on by all the psychologists, is called test standardization.

III - Criteria of a Good Test.

A) The reliability of the test:

It is the internal consistancy of a test or the accuracy with which it yields consistent results on testing and retesting. No test can be perfectly reliable because small or large errors creep into the results from different sources. The most common sources of inaccuracy or variation in performances of the individuals on a test may be summarized in the following:

1 - Actual differences among individuals in the general trait set to be measured.

2 - Specific abilities required in a particular test; or specific disabilities in the functions being tested.

3 - Skill in taking tests, being "test wise".

4 - The acquisition of a particular piece of knowledge or information required in a test by "chance".

5 - Effects of practice acquired by (previous test taking).

6 - Normal fluctuations in performance from time to time.

7 - The specific personal characteristics of the testee himself from the point of motivation, health, energy level, and emotional status.

8 - Physical conditions in the time of testing such as heat, light and ventilation.

9 - Accidental factors in the time of testing such as noise, interference, broken pencil, misunderstanding of instructions, etc.

10 - The successful guessing of some answers.(1)

Consequently evaluation of the reliability of a test is to find out how much of the variation in the individuals scores is the result of actual differences between them and how much is the result of the errors of measurement.

Before considering the methods of estimating the reliability of a test we are going to mention three general principles:

a) The reliability coefficient is affected by the length of the test.

b) The value of reliability coefficient changes with the scatter of scores in the group studied.

c) A test may give different reliability coefficients in different levels of ability.

---

(1) Freeman, op. cit. P. 17.

The importance of lengthening tests is that, with every question added the sample of performance becomes more and more adequate. But extreme increases in the test length, however, introduce effects of boredom and actually reduce reliability.

As for the second principle, when the spread of individuals' scores becomes smaller the same reliability coefficient becomes less significant because most of the differences in this case are due to errors of measurement. Thus "a test which has a satisfactory high reliability coefficient for use with groups containing wide differences in ability may be unsatisfactory for comparing people in a highly selected group."(1)

According to the third point, tests, no matter how reliable, give inaccurate measures when they are used for students of different level of ability. For example when it is given to students whose ability is less than those for whom it is attempted, the chance factor in the responses plays a greater role and the reliability of the test becomes less significant. Hence variability of scores may be different at different levels and variations due to errors of measurement may operate differently at each of the several levels.

Techniques have been devised for evaluating test reliability. Three procedures are the most common:

1. Two equivalent forms of the test are administered and the two sets of scores are correlated.
2. A single form of the test is given twice and the correlation between the two is computed.

---

(1) Cronbach, op. cit., P. 62.

3. The items of a single test are divided into two equivalent parts and the correlation between the two sets of scores is obtained.

Every procedure of these has a certain limitation. The first one requires more time. Moreover the two forms attempted to be used may be different in content thus diminishing the reliability of the test. On the other hand there must be a few days between testing and retesting in order to decrease the effect of experience got by the pupils in attempting the first form.

As for administering the same test twice, this procedure has some of the above mentioned disadvantages. There is also an opportunity before students to benefit from the recall of answers of some items in the first testing. But some psychologists refuse to accept this last statement because, as they state, the number of items is too large to allow the student to retain many of the answers especially when the time between testing and retesting is more than a week.

The third method is sometimes called the split half method in which the items of the whole test are divided into two equivalent parts in respect to:

a) mental processes included

b) difficulty

c) number of items in each part.

The subdivision is generally made by taking the odd-numbered items as one part of the test and the even-numbered items as the other. The score of every individual for each of the sub-

divisions is obtained and the coefficient of correlation between the two sets is computed. The result is at last corrected by a statistical formula (Spearman-Brown),[1] because the whole test is more reliable than each of the two subdivisions and the formula is intended to indicate the reliability of the entire test.

The coefficient of reliability by this procedure indicates how accurately the test measures the individual's performance for a given time. It shows to what degree the individual's score would change if more items representing the same functions are added. Therefore "the coefficient here does not describe entirely the reliability of the test which is to be used for periodical or predictive purpose. In this case retesting by using the same form is desirable."[2]

The coefficient of correlation obtained by this last procedure is higher than might be found by other methods since the two sets of scores in these methods are taken under the same conditions.

B) <u>The validity of the test</u>:

The validity of a test is defined as the degree to which the test measures what is attempted to be measured. And this cannot be if the test is not reliable. However, a test may be reliable without being valid. Besides, Greene comments,

---

(1) The formula is found in Chapter V, p.105

(2) Anastasi & Foley - Differential psychology, The Macmillan Company, New York, 1949, P. 44.

"Tests cannot be described as valid in general terms, but only in connection with their intended use and at the intended ability level of pupils."(1)

The criteria used for determining validity are different in kind and in number according to the purpose of the test and the availability of these criteria, e.g. in constructing intelligence tests some or all of the following criteria are usually used: scholastic marks, teachers' judgments of individuals' abilities, cumulative scholastic averages over a period of years, number of school grades completed, chronological age and known groups (gifted groups or mentally deficient groups).(2)

Often tests designers validate their tests by correlating them with other well-established and standardized tests. The test maker here must be very sure of the suitability of the test chosen as a criterion for this purpose.

The most frequently used technique in determining the validity of a test is to correlate test scores against each criterion. "Rarely the validity coefficient rises above .70."(3) "The maximum validity a test can have is the square root of the reliability coefficient."(4) However, "no coefficient of a given size can be specified as representing or not validity"(5) because

---

(1) Harry A. Greene, Albert N. Jorgensen & Raymond Gerberick, "Measurement & Evaluation in the Secondary School", Longmans, Green, New York, 1943, P. 61.

(2) Freeman, op. cit., P. 21

(3) Cronbach, op. cit., P. 58.

(4) Ibid., P. 73.

(5) Freeman, op. cit., P. 23.

this depends upon the trait or ability measured and upon the provided conditions. Therefore any coefficient which is positive has some value. Some coefficient of only +.30 are regarded as useful but most validity coefficients however should be well above this.

When a test has a low or moderate correlation with a single criterion, this correlation may increase when it is correlated with several criteria taken together as a composite. Consequently when the correlation found between the test and one criterion is low this result may still be regarded as useful since often more than two criteria are used. Thus the degree of correlation increases.

Moreover as the supply of men and the training conditions change from time to time, the test maker must revise his validity coefficient periodically.

There is a new tendency at present which suggests a new concept of validity called "factorial validity". It is the validity of a test for one type of ability only. Thus every trait represented in the test must have its own coefficient of validity.

C) <u>The validity of the items or item analysis</u>:

The validity depends not only upon the validity of content in general but also upon the validity of the individual items composing the test. These items play an essential role in determining reliability as well as validity. Therefore objective evidence concerning item validities is indispensable in the

standardization process. This evidence is secured only through the administration of the test in its preliminary form to a large number of individuals representing the population for whom the test is attempted.

Three aspects of the items are usually studied:

1) The level of difficulty of each item.
2) The correlation of each item with the subtest score of which it is a part.
3) The degree to which each item descriminates between a high group and a low group.

1) Modern practice in the arrangement of test items is to present items covering a wide range of difficulty in ascending order from the very easy to the most difficult. This arrangement provides for every person to work on items within his level of mastery without being discouraged with difficult items outside his abilities.

The difficulty of a test item is expressed indirectly in terms of the percentage of individuals of a certain group who are able to pass it. The majority of test specialists prefer to select a few easy and a few difficult items but to have the majority near the 50 per cent difficulty level. "It is desirable however to have at least one item that can be passed by all and one item which no one can pass."(1)

2) The second of the previous aspects, is to correlate

---

(1) Freeman, op. cit., P. 25.

each item against the score of the subtest of which it is a part for example, vocabulary, arithmetical problems, etc. in order to find out whether the performance on this item is consistent with the performance of the subtest as a whole. Because the assumption here is that all the items in a subtest measure the same psychological processes.

3) The basic function of all measurement is to place individuals along a scale of ability or achievement. Such function implies discriminative power on the part of the test. Consequently each item should descriminate between two extreme groups the highest and the lowest. This classification of the individuals under two groups is based upon scores on the test as a whole, or upon external criteria. Some test makers take only the highest and lowest ten percent, others prefer to take the highest and lowest fourth of all the individuals taking the test or the upper 50% and the lower 50%. Kelley's procedure is the highest and lowest 27%.(1) Another criteria for the validation of items is that the percentage of correct responses should be significantly greater in any particular grade than in the grade below.(2)

Thus, selecting the items of the test according to some of these criteria, the test designer makes them up at last into a test.

---

(1) Ibid, P. 26. (For the implications of this point see Chapter V, P. 38 of this paper.)

(2) G.C. Scott, Measuring Sudanese Intelligence, The British Journal of Educational Psychology, Vol. XX, Part I, February 1950, P. 44.

D) The fourth criterion of a standardized test is objectivity which makes for the elimination of opinion, bias or judgment in administering or scoring the test so as to provide a basis for uniformity of interpretation of results. A convenient form of answer key is always provided. Besides, there is always a manual of directions, based on an elaborate experimental basis and carries usually complete instructions for administering and scoring of the test which must be followed exactly by the tester.

E) Every standardized test is also accompanied in the test manual or elsewhere by adequate tables of norms adapted in type to the age and grade levels for which the test is intended. The norms are usually in form of age, grade, percentile rank, decile rank or standard score.

Norms may be defined as average scores determined by actual measurement of a large group of persons who are representative of a specified population. Therefore they are dependent upon the conditions and opportunities available to these persons during the course of their development. "And since methods, cirruculums, administrative practices, etc. are likely to change with time, norms for old unrevised tests should periodically be brought up to date to be adequate for use."(1)

Whether norms are satisfactory for comparability depends on three things:

---

(1) Herbert S. Conrad, "Norms", Encyclopedia of Educational Research, ed. Walter S. Monroe, 1941.

1 - They must be based on a sufficiently large group. Because if the number of cases is limited, the norms may be inaccurate due to accidental selection of good or poor cases.

2 - The standard group must be representative: representativeness is more important than the first condition. Especially if the large number of cases is composed of persons who are not typical of those for whom the test is intended.

3 - Norms are applicable only to samples and individuals who resemble closely the typical case of the standard group. Unfortunately publishers and designers of tests rarely describe in detail the standard group in order to enable the tester to evaluate the degree of conformity between the standard and his group.

"Proper detailed interpretation of test scores clearly rests not merely on the conscientious application of norms but also upon the intelligent evaluation of all identifiable factors likely to cause departure from the published norms."(1)

Cronbach also adds, "Before we compare a pupil with a set of norms, we must consider the educational policies of the schools where the norms were obtained."(2)

IV - Intelligence & Intelligence Tests.

Intelligence tests are an essential branch of the measurement movement in the West around which many debates and

---

(1) Herbert S. Conrad, Encyclopedia of Educational Research, ibid.

(2) Cronbach, op. cit., P. 77.

discussions have arisen. Intelligence is not a simple entity which can be measured directly. But it is manifested only through the individual's behavior which is always conditioned by heredity as well as by previous experience or training.

A) Definitions of Intelligence

Many definitions have been tried by psychologists in order to measure and assess intelligence. Here are some of them:

Binet for example defines it as "the capacity to reason well, to judge well and to be self-critical."

Terman in 1921 stated "an individual is intelligent in proportion as he is able to carry on abstract thinking."(1)

Bagley says, "General intelligence is the ability a) to adapt oneself through a process of judgment to new situations, b) to draw general inferences from specific experiences, c) to learn rapidly."(1)

Goddard reaches the conclusion that "intelligence is the degree of availability of one's experience for the solution of immediate problems and the anticipation of future ones."(2)

Generally definitions of intelligence range all the way from the broad, inclusive description of intelligence as the ability to make adaptations to a changing environment to the more specific description of intelligence by Spearman as the education of relations and the education of correlates.

---

(1) Goddard, "What is Intelligence", The Journal of Social Psychology, Vol. 24, 1946, P. 51.

(2) Ibid, P. 69.

Following Freeman in his classification the different definitions of intelligence can be grouped into several categories:(1)

1 - One group of definitions emphasizes adjustment or adaptation of the individual to his environment or changing conditions. According to this group intelligence is a general mental adaptability to new problems of life.

2 - The second group states that intelligence is the ability to learn and the more intelligent the individual the more he is able to learn and to get experiences.

3 - Others find intelligence as the ability to carry on abstract thinking or the effective use of concepts and symbols in handling problems.

Studying these groups of definitions thoroughly, we find that no type of them can be separated from the other. Learning capacity is essential for adjustment and adaptation although other non-intellectual factors affect adjustment. Though learning and previous experiences are indispensable for adjustment yet learning in the sense of repeating the facts without insight is not enough to describe intelligence. Therefore the more intelligent the person the more he can reorganize his past experience to fit the novel situations. Also the ability to use symbols and concept is itself the result of learning and development. The child cannot use symbols in the beginning but once he begins to use them they promote his learning and consequently they contribute to his ability to adapt to

---

(1) Freeman, op. cit., PP. 68-70.

new situations by extending his horizon and enlarging his range of behavior.

In recent years, however, other new definitions of intelligence have appeared. They combine and extend the three previous groups of definitions and include other new aspects of intelligence which are not intellectual. Stoddard's, for example, is one of these definitions. He states, "Intelligence is the ability to undertake activities that are characterized by 1) difficulty, 2) complexity, 3) abstraction, 4) economy, 5) adaptiveness to a goal, 6) social value, 7) the emergence of originals and to maintain such activities under conditions that demand a concentration of energy and a resistence to emotional forces."(1)

This definition expands the three above mentioned definitions and adds more to the clarification of intelligence but the inclusion of "social value" in a definition of intelligence is of doubtful validity because there is no absolute static social standard. Social value is something subjective therefore it cannot be measured objectively. Moreover can we consider those geniuses who live out of harmony and do not conform to their social environment, as not intelligent? According to Thorndike these kinds of people may be lacking in social intelligence while they may have high abstract or concrete intelligence. Intelligence, as stated by him, functions in three main spheres:

---

(1) Stoddard, "The Meaning of Intelligence", The Macmillan Company, New York, 1947, P. 3.

a) the concrete which is the ability to understand and deal with things,

b) the social or ability to understand and deal with persons,

c) the abstract or ability to understand and deal with verbal and mathematical symbols.

The significant thing about this classification lies in its indication that what an individual attains in an intelligence examination depends to a considerable degree on the type of test used. This definition thus implies the devising of separate and specialized tests in each of these three aspects. But the inclusion of "social value" in the intelligence tests is hardly present in current tests; although, of course, some psychologists like Stoddard, declare that it should be. The third type (the abstract) is the one aspect of intelligence that receives greatest weight. T.L. Kelley, has shown that ordinarily "about 90 percent of the general intelligence tests and the all-round school achievement tests measure the same functions."(1) Others add "Most intelligence tests may be regarded as measures of scholastic aptitudes, or ability to succeed in our schools."(2)

The verbal loading of the current tests of intelligence is explained by Anastasi and Foley on cultural grounds. According to them, "Intelligence can be defined only with reference to a particular setting or environmental milieu. This view

---

(1) H. Garrett and Schneck, "Psychological Tests, Methods and Results", Harper and Brothers, New York, 1933, Part II, P. 5.

(2) Anastasi and Foley, op. cit., P. 487.

immediately suggests that there are, not one, but many definitions of intelligence. Within our cultural setting, intelligence apparently consists in large part of verbal ability."(1)

The last two conditions in Stoddard definition "concentration of energy" and "resistance to emotional forces" though they affect the individual's effectiveness, their introduction in a mental test may confuse the results and invalidate them. However we cannot give a precise judgment on what the research about this point in future will reveal but at present the introduction of non-intellectual aspects in mental tests, is not yet practiced. Nevertheless the actual intelligence tests measure motivation and concentration of energy indirectly, the examiner tries to provide the conditions where the testees do their best and if an individual is emotionally blocked or not well motivated or handicapped by certain external factors the trained tester can notice this clearly in an individual test. But during group testing there may be instances of such persons without that fact being known to the examiner. This can, however, be overcome by scrutinizing and evaluating each one's rating in the light of other and perhaps conflicting evidence. Therefore a variety of personality rating scales and inventories and projective techniques have been constructed to evaluate these non-intellectual traits.

David Wechsler includes also those non-intellectual factors such as drive and incentive in his definition: "intelligence

---

(1) Anastasi and Foley, op. cit., P. 488.

is the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment."(1)

The new aspect in this definition is the author's declaration that the individual's intelligence is revealed or characterizes the individual's behavior as a whole.

There are still some psychologists who try to define intelligence operationally like Gates. They think it is irrelevant or misleading to define intelligence. Therefore they say, "intelligence is whatever the intelligence tests measure."

This may be good now but it does not help persons who start out to build tests. Quoting Stoddard, "At that time a test measures nothing. What does the test constructor put into his tests and why?"(2)

Finally, surveying the different group tests constructed from the different definitions of intelligence, one finds that though they have some differences in their content, nevertheless they have many things in common. This implies that they are interdependent and their differences are largely in respect of emphasis. It is shown also that most of the tests do not incorporate all aspects of the comprehensive definitions of intelligence such as those of Stoddard and Wechsler.

B) <u>The nature of intelligence</u>:

The nature of intelligence has occupied the attention of psychologists ever since they began to measure it. Thus they

(1) David Wechsler, "The Measurement of Adult Intelligence", the Williams & Wilkins Company, Baltimore 1944, P. 3.
(2) Stoddard, op. cit., P. 3.

started with long and intricate procedures of analysis, known as factor analysis in order to discover the elements or components of intelligence.

The principle on which the experimental method in determining the nature of intelligence is based can be summarized as follows: a large group of tests different in character are given to a large and adequate samples of population. Then the results obtained from every test are correlated with the results of the other tests. The coefficients of correlation thus obtained are subjected to long and various statistical techniques. As in all sciences, the particular theory resulting from this statistical analysis depends upon the analyst's own interpretation. Therefore the specialists in this field have not agreed upon the reality of mental factors. Cronbach finds that this is due to the misunderstanding of the statistical methods used for this purpose. The factor analyst, according to him, is not like the chemist in his search for elements. The latter has only one answer to his question while the former has many answers all equally true but their convenience is not equal.(1) Then Cronbach compares the factor analyst with a photographer who tries to picture a building "whenever he sets his camera he will lose some information, but by a skillful choice he will be able to show a large number of important features of the building."(2) But theoretical differences among psychologists concerning the nature

(1) Cronbach, op. cit., P. 208.

(2) Ibid., P. 208.

of intelligence, have not affected the kinds of tests resulting to any great extent. There are three main theories concerning the nature of intelligence:

1) The Multi-factor theory of Thorndike:

It is at one extreme of these theories. According to it, intelligence consists of a multitude of independent factors and every mental act, involves a number of these minute elements functioning together. As stated by this theory the correlation between two mental tasks depends upon the number of common elements involved in the two acts. Thorndike adds however, that some mental activities have so many of these elements in common, that they can be grouped under special names (such as verbal meaning, arithmetical reasoning, etc...) This is shown apparently in his test CAVD devised to test abstract intelligence, which is composed of four parts: sentence completion (c), arithmetical reasoning (A), vocabulary (V), and following directions (D). Thus practically the theory of the multitude of minute elements does not differ in the content of the test from other theories as will be seen later.

2) The Bi-factor theory of Spearman:

It is at the opposite extreme of Thorndike's theory. According to it, all intellectual activity is dependent essentially upon two factors, one general (g) common to all mental activities and the other a "specific factor"(s) peculiar to each particular ability. This statement may be represented by a mathematical form by showing the score (S) of an individual on a

given mental test by the equation $S = a_1 g + a_2 s, \ldots$ where the letters "a_1" , "a_2". represent the loading of the two factors "g", "s" respectively. The g factor is constant for every individual although it varies greatly for different individuals. Moreover it enters in all mental activities though in varying amounts according to the demand of the mental task upon general intelligence.

As stated by Spearman this factor "g" has to do with Speed and Clearness but it is independent from row learning and memorizing, "It was found to be equally concerned with each of the two general dimensions of ability, Clearness and Speed ... But it revealed a surprisingly complete independence of all manifestations of Retentivity."(1)

In explaining the correlation between intelligence factors (g) and (s), Spearman says, "Under all conditions (s) is wholly uncorrelated with (g). Further with certain exceptions every (s) is independent of every other one, unless the two operations are closely similar."(2)

Later researches led Spearman and his supporters to conclude that there are certain "group factors" which enter also in some mental activities, quoting him, "Among the exceptional cases where, on the contrary specific correlations and group factors do become of appreciable magnitude, the four most important have been in respect of what may be called the logical,

---

(1) Spearman, "The Abilities of Man", Macmillan & Co., London, 1927, P. 412.

(2) Ibid., P. 413.

the mechanical, the psychological, and the arithmetical abilities. Each of these group factors has been discovered of sufficient breadth and degree."(1)

Hence, all mental activities, according to this theory, have in common some of general factor, their own specific factors and each of these activities might be also a member of activities that have their own group factors. The general factor, however, still remains the primary one and it is the essential measure of intelligence.

But what psychological entity is to be identified by (g)? and what is the nature of (g)? Spearman discusses many possible answers to this question. He rejects such terms as "adaptation", "abstraction", "attention" and "will" partly because they are ambiguous and partly because they represent entire mental operations. In Spearman's own words, "Such terms as adaptation, observation, and so forth, denote entire mental operations; whereas our (g), as we have seen, measures only a factor in any operation, not the whole of it."(2)

Spearman hesitates to identify "g" with intelligence because of the wide different names given to the latter term. "Such a factor as this can scarcely be given the title of 'intelligence' at all, being evoked to explain the correlations that exist between even the most diverse sorts of cognitive performance, it does not deserve a name appropriate to any one

---

(1) Spearman, "The Abilities of Man", Ibid, PP. 241, 242.

(2) Ibid., P. 88.

particular sort."[1]

Consequently he regards "g" as "energy" which serves the whole cortex which can be transferred from one mental operation to another different one. Thus he likens "special ability" to one of a number of machines in a factory, and "general ability" to the energy which operates the machines when they are performing their functions. We quote him, "For this factor (s) also a physiological substrate has been suggested, namely the particular group of neurons especially serving the particular kind of operations. These neural groups would thus function as alternative "engines" into which the common supply of 'energy" could be alternatively distributed."[1] Some psychologists do not agree upon such an analogy which differentiates between (s) and (g) in terms of kind. We quote Hamley in this point, "Such an analogy, however, scarcely fits the facts, for it draws a distinction in kind between the specific and general factors. It is doubtful, however whether an analogy as simple as the above exists."[2]

Wechsler believes moreover that Spearman considers (g) & general intelligence as being equivalent and he continues to say that the Spearman's mathematical relationship of "g" and (s)[3] proves his statement. For by pooling such tests, the

---

(1) Spearman, "The Nature of Intelligence and the Principles of Cognition", MacMillan & Co., London, 1927, P. 5.

(2) H.R. Hamley, "Intelligence Testing", the Year Book of Education, Evans Brothers Ltd., London, 1935, P. 424.

(3) The equation is on page 38 of this chapter.

"g" being common becomes cumulative, whereas the specific factors tend to cancel each other. In Wechsler's opinion such scale would not be a good measure of general intelligence "because it would eliminate a number of abilities essential for effective behavior ... individuals attaining identical scores in intelligence tests cannot always be classified in the same way."(1)

Briefly the nature of (g) and the extent to which it can be identified with general intelligence is still a debatable and speculative matter.

Some configurationists have seen that Spearman's theory is equally answerable in the whole - part brand of interpretation. Thus Wheeler says: "every form of behavior is a pattern response, the most obvious features of which are its specialized aspects, but the total pattern is always functioning. Whatever the wrist is doing at any time is part of something the arm is doing."(2)

In every mental activity, in term of Gestalt psychology, there is a total intellectual pattern which is the aspect of insight of the individual's total mental life, it is represented by the arm in the illustration. The Spearman's S's as stated by this theory are not factors separate from (g) but they are differentiations of it. Going back to the arm-wrist illustration we find that when the wrist is bent there is also an arm movement and whenever the arm is moved, the wrist is involved.

Hence the configurationists come to the conclusion that

---

(1) Wechsler, The Measurement of Adult Intelligence, op. cit, P. 8.

(2) George W. Hartmann, "Gestalt Psychology" (A survey of facts and principles)" Ronald Press Company, New York, 1935, P. 265.

there is no reason to believe that the person cannot reach the same degree of maturation in every activity which he is engaged in. The brilliant individual, they claim, must be brilliant in every thing he does, provided the same opportunities, interest and proper stimulation exist.

3. Thurstone and the Group-factor Theory:

This theory is intermediate between that of Spearman's and Thorndike's. All mental operations, according to this theory, may be looked upon as groups and every group of them has in common a "primary" factor which characterizes it with psychological and functional unity and which differentiates it from other mental operations having their own unifying "primary" factor. Each of these primary factors is said to be independent from one another. Thurstone and his collaborators concluded that there are six primary factors: (1) the number factor (N), (2) the verbal factor (V), (s) the spatial factor (S), (4) the word fluency factor (W), (5) the reasoning factor (R) (induction and deduction are involved) and (6) the Rote Memory factor. There are still two other possible factors (7) Perceptual factor (P), and (8) Deductive factor (D).

One important factor found in studies other than Thurstone's is a speed factor. This factor was held constant in Thurstone's studies. Indeed the proponents of this theory do not claim that the exact number of these primary factors is known. Frederick Davis states, "the results of testing hundreds of thousands of men in the armed forces and of analyzing these

data suggested to many psychologists that the number of basic mental abilities may often have been under-estimated."[1] He doubted also that the names attached to these factors accurately describe their nature.

Thurstone supposes that the individual's performance on a test or his ability can be broken down into a certain number of factors, x_1, x_2, x_3 ... and by the help of the statistical analysis the most important factors can be determined. Thus the individual's ability or his score (S) on test (A) or test (b) can be represented by the equations:

$$S_a = a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots$$

$$S_b = b_1 x_1 + b_2 x_2 + b_3 x_3 + \ldots$$

in which a_1, a_2, a_3, b_1, b_2, b_3 represent the factor loadings.

Theoretically the coefficients of correlation among these factors must be zero but actually it has been found that they are positively intercorrelated. This means that primary factors are divisible and that other factor or factors must be at work with them. In other words, the authors of the test either have not been able to sample them purely or this is impossible to be devised. Consequently Thurstone has introduced to his primary factors a general factor which has a second place, whereas Spearman looks at it as the first. The difference between the two is a matter of emphasis.

Two consequences of group factor analysis have been stated:

---

(1) Cronbach, op. cit., PP. 210, 211.

a) It tends to clarify the nature of the items included in a test so that it helps the test maker in a more definite and purposeful selection of test materials.

b) The tests based on this hypothesis usually yield a mental profile for every person and not a single index and this is much more useful to understand the pupils better. But though this requires an insight by the examiner yet Thurstone declares that, "There is no question but that the profile is more helpful than IQ in the interpretation of educational and behavior problem."(1)

Freeman finds a weakness of the group factor type of tests and this is, "they ignore the fact that intelligence expresses itself in behavior as a combination, a unity of functions, not as a series of independent factors."(2)

The test maker usually chooses before beginning his test what factor (bi-factor or multiple-factor) he wants. Any type of factor solution can be transformed to another by certain mathematical methods. "For set of positive intercorrelates neither factor description is more "right" than the other. The difference between them is one of emphasis and convenience."(3)

It is true that tests based upon Thurstone's and Spearman's hypothesis are differently organized yet the content is not basically varied. The differences at present about the

---

(1) Ibid, P. 209.

(2) Freeman, op. cit., P. 88.

(3) Cronbach, op. cit., P. 198.

nature of intelligence is mostly limited in theoretical matters. Therefore intelligence is like electricity. It is measured though its nature is not yet agreed upon. "In terms of dynamics, there would exist no possibility of measuring a native capacity, or potential, if one existed, the only kind of energy that can be measured is kinetic energy, or work. In human life we call it achievement."(1)

C) The Types of Intelligence Tests & Their Different Characteristics:

Current intelligence tests may be classified either according to the content such as verbal and non-verbal, or performance tests, or according to their administration, such as group and individual tests.

Verbal tests are the most commonly used tests. They are entirely verbal in character. They contain symbolic materials (language and number).

While in Performance or non-verbal tests the role of language is eliminated from their content or responses although generally this condition is not included in the directions. The test materials of a performance scale are concrete objects, pictures, mazes, and geometric figures. They must be administered individually or in small groups while group non-verbal intelligence tests may be administered to a group of persons - large or small - all at one time.

Performance and group non-verbal intelligence tests

---

(1) Raymond H. Wheeler, The Science of Psychology, Thomas Y. Crowell Company, New York, 1940, P. 171.

are usually used with persons who have language handicap due to foreign background, deafness, and with illiterate or non-English speaking adults. They are also valuable to diagnose well those cases who get .60 or .75 on verbal tests.

Since the use of language (vocabulary, sentence completion, analogy, proverbs, etc. and number series) demand much more maturation than the immediate concrete situation, psychologists find that non-verbal tests are more useful with younger and inferior or handicapped subjects, than superior and older individuals. And though the coefficients of correlations between verbal and non-verbal tests are significant, yet they are not equivalent and do not involve the same mental functions. This suggests to us to use the two types, verbal and non-verbal as supplementary rather than substitute or equivalent.

On the other hand when we classify intelligence tests under individual and group tests we find that individual intelligence tests are usually given to one person at one time. They form actually the most accurate device for measuring intelligence. The examiner here, on the contrary to group tests, does not merely rely upon the quantity of score obtained but he observes each reaction made by the subject. While group intelligence tests are less time consuming and quickly administered and scored, nevertheless they are less informative than individual tests. But under favorable conditions, group tests are highly predictive and as reliable as individual tests.

D. The Use of Intelligence Tests:

Intelligence tests have been applied for many problems, educational, sociological and psychological.

a) In schools they are used widely:

1) for selection of college students;

2) to locate the mentally retarded and accelerated children so as to subject them to appropriate teaching methods and materials;

3) In high schools they are useful in helping students to select their courses and in giving vocational advice, without indicating the specific vocation a person should choose;

4) for common educational investigation and survey, such as correlations between intelligence quotients and scholastic progress, increase in average level of IQ from grade to grade, differences between pupils in various schools within the same community and the effects of different methods of instruction upon pupils' mental ability.

b) For general and psychological problems such as individuals' differences in relation to sex, race and nation, family similiarities and the inheritance of intelligence, effects of changed environment upon mental level, the nature and course of mental development, the nature of intelligence and the constancy of IQ.

c) They are also used in solving many sociological problems:

1) they are being employed more and more in Juvenile courts. For example a bright maladjusted youngster is quite a different problem than a moron.

2) in war, for classifying men for different branches of the service.

3) Intelligence Tests & to What Extent do They Measure Intelligence?

It was said sometimes that intelligence tests do not differ from achievement tests as both are based on factual data and learned skills. This is right to a very limited degree. But the factual materials in the intelligence tests are supposed to have been available for all persons except those who are living in a very exceptional environment. Consequently these tests remain mainly measures of persons' abilities to see relationships, to draw inferences, to manipulate, to compare, to solve mathematical problems and to contrast those commonly known facts. To identify these mental activities with general intelligence is still doubtful. Intelligence is all these and yet something more. For example current intelligence tests do not measure judgment and wisdom that reflect the effect of experience nor productive thinking or intellectual originality. In general as the individual grows older the less the intelligence tests, available at present, measure effectively his global intelligence.

Indeed any test measures only a part of the total personality of the individual. This part is an integrated division of the whole. Therefore it is influenced by it. Hence, whenever there is a necessity to understand the whole individual, various tests of general and specific abilities and personality scales are used to supplement one another.

CHAPTER III

MENTAL TESTING IN THE ARAB WORLD

Testing movement in the Arab world started approximately twenty-five years ago and few adapted tests appeared in Egypt, Iraq and Palestine. The aim of these tests was different in the different countries and generally the sound scientific basis which makes of a test a standardized one was lacking. Moreover no better tests have appeared at present in the Arab world to show a new trend in this field or to reflect the degree of improvement which may have been attained. But unfortunately this field showed more enthusiasm in the beginning than it does today.

However to give the reader a more clear picture about this movement in the Arab world we are going to describe briefly according to the information available in our hands, some of the different tests which have been made up till the present day in this part of the world.

I. The Tests Adapted in Egypt.

Intelligence tests were not known in Egypt before 1928 when the Government invited two swiss experts to study the educational system in Egypt so as to suggest the improvements which must be introduced to that system. Claparede who was one of these two experts preferred before giving any suggestion to

study the conditions scientifically. Therefore he chose a series of tests, group and individual, and after adaptation he gave them to students in different schools, analyzed the results and interpreted them in a report presented to the Government. One of the tests used was that of Ballard. It was an intelligence test which was prepared in England for grades 8 and 12. It consisted of 100 items presented to the students orally. Mr. Ismail Al-Qabbani translated the test into Arabic and has adapted it to Egyptian environment. Then it was given to 3000 boys and girls in the primary, elementary, secondary, industrial and in the elementary teachers college. Finally the curve of Egyptian students was compared with that of the Swiss and English students.

After Claparede, Ismail Al-Qabbani followed up the study and made many modifications in it. He studied again the items of the test and eliminated those which appeared to be unsuitable to Egyptian students. The selected items became 46 administered in two periods of forty-five minutes each. He introduced also some changes in choosing the sample. Then he computed the reliability and it was found to be between .86 and .90.

To measure the validity of the test he correlated the results of his test with the students' grades in the school and with their teachers' estimates. The latter proved to be unreliable while for the former he gave a chart as it is shown below in Fig. (1). In this figure A, B, and C represent respectively the superior, the average and the inferior students

according to their grades in the school. The test results indicate for example in A that only 69 percent of the students considered as superior in school, are superior while 22 percent of them are average students and 9 percent are inferior; in B

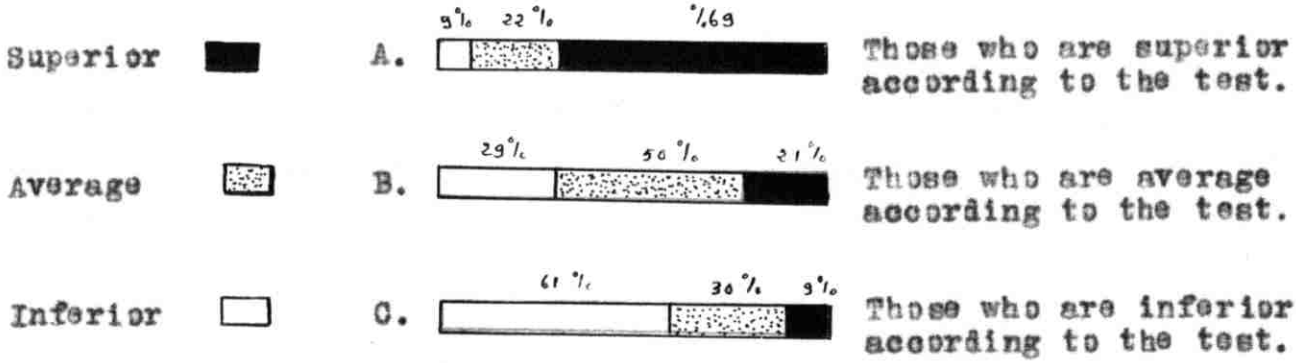


Fig.(1) The correlation between Mr. Al-Qabbani's test results and the students rank in the school.

the test shows that only 50 percent of the group, considered as average in the school, are average students and in C only 61 per cent of the inferior group are really inferior while 30 percent of them are average students and the rest 9 percent are superior.

In 1938 Ismail Al-Qabbani published these results got in the elementary schools in Cairo in a book. The test was called, "Ikhtibar Al-Zaka' al-Ibtida'i",(1) but as yet, no extensive enough sample of this new adapted test has been used, in order to make it a standardized one.

---

(1) Ismail al-Qabbani, "Ikhtibar al-Zaka" al-Ibtida'i", Lijnat Al-Ta'lif wa-Al-Tarjama, Cairo, 1938

اسماعيل القباني ـ اختبار الذكاء الابتدائي لجنة التأليف والترجمة بالقاهرة سنة ١٩٣٨

(Notice) (Al-Zaka' is better pronounced 'Al-dhaka' according to phonetics).

Before Mr. Al-Qabbani's book another one had been published about this subject by Dr. Hasan 'Umar and it was called "Kiias Al-Zaka'" or "Testing of intelligence" and it contained an adaptation of Standford-Binet revision (1916); the questions were put in the colloquial language with the necessary pictures so as to be understood by children.

Many other Egyptians such as Dr. 'Abed al'Aziz Al-Kūsi, Mr. Mouhammad Fuad Jalal, Mr. Kamel Al Nahhas, have worked on a number of intelligence tests.

Here are some of the other adapted tests in Arabic that have appeared in Egypt:

١ ــ مقياس "استنفرد بينة "الذكاء ــ حسن عمر

1. "Mukias Stanford Binet Lil Zaka'" by Hasan 'Umar(1)

٢ ــ الذكاء المصور

2. "Al-Zaka' al-Mūsawwar" (non-verbal test)(1)

٣ ــ اختبار الذكاء الثانوى ــ اسماعيل القباني

3. "Ikhtibar Al-Zaka' Al-Thanawi" by Ismail Al-Qabbani(1)

٤ ــ اختبار الذكاء المتوسط ــ اسماعيل القباني وعبد العزيز القوصي (غير لفظي)

4. "Ikhtibar Al-Zaka' Al-Mūtawassit" by Ismail Al-Qabbani(1) and Abdul'Aziz Al-Qūsi (non-verbal)

5. "Ikhtibarat Spearman Al-Hūssiia Lil Zaka'" by Ismail Al-Qabbani and Abdul'Aziz Al-Qūsi (Non-verbal test.(1) The manual is also non-verbal)

٥ ــ اختبارات سبيرمان الحسية للذكاء ــ اسماعيل القباني وعبد العزيز القوصي
(كراسة التعليمات غير لفظية)

---

(1) Lijnat Al-Ta'lif Wa Al Tarjama, Cairo, Egypt. (The years of publication are not mentioned here because the manuals were not available).

6. "Ildtibar Al-Zāmlek Lil Zaka'" by Muhammad Kamel Al-Nahhas (for high schools).
اختبار الزمالك للذكاء ـ محمد كامل النحاس ـ (لفظي) يمكن استعماله للمدارس المتوسطة

There are some tables of averages for some of these tests based on relatively small samples. But most of them are still in the try-out process hence many modifications and experiments are needed so as to become standardized and suitable for use.(1)

This is a short summary of the tests and studies performed in Egypt during the last twenty years. Not one of these tests is officially used in the schools. For, since the educational systems in the Arab countries are centralized, no government has yet included these new techniques of measurement in their educational projects. The writer could not get any information about their validity or reliability because nothing has yet been published about them. In our opinion they are still in their very preliminary stages. Thus we cannot evaluate them justly at present.

II. Mental Testing in Iraq

The movement of Mental testing in Iraq began twenty years ago and the early years of this movement were relatively the flourishing years. Some of the tests were translated into

---

(1) 'Atiya Mahmūd Hānā, "Al-Zaka' wa Makayisuhā", Al-Nahdā Al-Masriia, Cairo, Egypt, 1949 (the introduction).
(١) عطيه محمود هنا "الذكاء ومقاييسه" ـ مكتبة النهضة المصرية سنة ١٩٤٩ ـ المقدمة.

Arabic and some were applied on some schools under the supervision of Sati' Bey Al Husseri. The aim of these tests was primarily to give the students in the teachers' colleges some examples of the new trend in evaluating students' abilities and aptitudes and to stimulate the attention of the teachers and those who are interested in education to this movement in the west. Here are the tests which were worked on in Iraq:

1. Translation & description of Terman Test (1)

2. A summary of Binet scale (2)

3. Army Alpha examination

It was adapted in 1929 and given to 320 students in the higher Teachers College and Elementary Teachers College, to students in the secondary, elementary and primary schools. (3) Then the comparison of the Iraquian students' results with those in America showed, according to the authors comments, that the students in Baghdad were not lower than American students in their intelligence. (4) Nothing else has been mentioned in the report about the test and its results.

4. "Al-Ikhtibar Al'Akli Al-Jam'i bil As'ila Al-Sharahyia" - Ballard. (5)

It consisted of 100 items and it could easily be applied in the elementary schools. Thus two elementary schools in Baghdad were selected and the sample was composed of 183

١) اختبار ترمن قائم على اختبار بينة
٢) مقياس بينة
٣) اختبارات ج. أ.
٤) الاختبار العقلي الجمعي بالأسئلة الشفهية - بلارد

(1) Majalat Al Tarbiah wa al Ta'lim Vol. 4,5,6, April & July 1928.
(2) Ibid, Vol. 6, 1928. (3) Ibid, Vol. 18 June 1929 P. 203.
(4) Ibid, (Vol.18 June 1929). P.308. (5) Ibid, (Vol.18 June 1929), P.309.

pupils. Again a comparison between the median of Iraquian students and the median found in the original Manual was the only computation made for the test.[1]

5. "Ikhtibar Th. M." (Thurstone-Mira)[2] ٥ـ اختبار ث.م. (ثرستون ـ ميرا)

It was composed of 70 items. At the end of 1928 the test was applied on 1029 students coming from different areas in Iraq to Baghdad for public examinations. The subjects of the sample were from different schools. 660 students from elementary schools, 87 from secondary, 253 from the intermediate and 32 girls from the Teachers College. Then a table showing the percentage of the different classes in every quartile was made. Two other tables were drawn also indicating the percentage of the scores of the different classes in the first half hour and in the second half of the test administration. But nothing about the reliability, validity or items analysis has been tried for the adapted test.[3]

There are still a small number of non-verbal intelligence tests which have been translated and sometimes applied. Here is a summary of them:

1. "Ikhtibar - Decroley - Tartib (As-suwar)"non-verbal"[4]

١ ـ اختبار دكرولي ـ ترتيب الصور

---

(1) Ibid,(Vol. 19 and 20, 1929),P. 527.

(2) Ibid,(Vol. 19 and 20, 1929),P. 384.

(3) Ibid,(Vol. 19 and 20, 1929),P. 400.

(4) Ibid,(Vol. 18, 1929),P. 358.

It was given to 48 boys in the first grade in the secondary school and 375 girls in the Teachers College. Then a table showing the percentage of the scores of the girls and boys was prepared with the comments that girls proved to be higher in this test than boys.(1) But what was the nature of the sample used? On what grounds were girls and boys chosen? Was the test valid? Was it reliable? Nothing has been done about all these questions. Therefore the reader wonders on what basis the conclusion was drawn.

2. "Ikhtibar Tamyiz Al-Sakhafat"(2) ٢ـ اختبار تمييز السخافات (في صور وعبارات)
(For descriminating absurdities - Semi performance test).
The sample was 374 boys in the Secondary School and Teachers College. Then a table was prepared to show the percentage of the right answers in the two groups without any comment or interpretation.(3)

3. There is still another test which was translated without application and it was called "Ikhtibar Al-Zaka' bi-Wasitat al-Rusum ٣ ـ اختبار الذكاء بواسطة الرسوم
It was originally made by Ballard for children 6 years old.(4)

Group Achievement Tests

1. "Mukyas Monroe Lil Hisab" (Arithmetic)(5) ١ـ مقياس مونرو للحساب
It was translated but no information is available whether it

(1) Ibid,(Vol. 19 and 20), 1929, P. 524.
(2) Ibid,(Vol. 18, 1929), P. 354.
(3) Ibid,(Vol. 19 and 20), 1929, P. 525.
(4) Ibid,(Vol. 19 and 20), 1929),P. 509.
(5) Ibid,(Vol. 18, 1929), P.338.

was applied in Iraq or not.

٢ــ مقياس كلاباريد للحساب

2. "Mukyas Claparede Lil Hisab" (Arithmetic)

The same exercises which Claparede used in Switzerland were used in Baghdad. The sample was composed of 972 students from the elementary, secondary and Teachers College. The results were classified according to grade and not age (because the latter, according to the author's comment, was not accurately known in Iraq). Then the results were compared to the norms found in Claparede's norms tables. At the end very brief comments were made about the weakness and strength of the different Iraquian students as compared with the Swiss students.(1)

Many other achievement tests in arithmetic were translated with the original norms tables but it is not known whether they were applied or not. These tests are as follows:

Mūkyas Courtis (2) مقياس كورتيس

Mūkyas Cleveland (2) مقياس كليفلاند

Mūkyas Ballard (2) مقياس باللارد

Mūkyas Al Sin (China) (2) مقياس الصين

There is also a test of morals the source of which we do not know. مقياس الشعور الاخلاقي Nothing is available about it except a description of the items and the stories included (3) in it.

---

(1) Ibid, (Vol. 18, 1929), PP. 338-342.

(2) Ibid, (Vol. 19 and 20), 1929, PP. 529 ...

(3) Ibid, (Vol. 19 and 20), 1929), P. 517.

After the period 1928-1929 no effort has been made to improve these tests, and no trial has been given in this field till 1946 when Mr. Muhammad Kamel Al Nahhas, the Egyptian test worker, tried to adapt Stanford Binet test (1916) which had been adapted to Egyptian colloquial language, to Iraquian children. But unfortunately no statistical results are available now to help us to evaluate this study. But what is known according to the report sent by the Iraquian government to the Education department in the A.U.B. in 1951, is that this test has not yet been standardized on Iraquian children, and the study is not going on today.

From the previous summary about the Mental Testing in Iraq one gets the idea that indeed none of these tests has been scientifically approached and interpreted. In none of them was the sample scientifically formed, even the validity, reliability or the item analysis processes were attempted. But remembering the time these studies were made (1928-1929) and the lack of experts in this field, we may say that as a starting step it was good if followed by better and more deeply studied tests, but unfortunately this preliminary movement died in its beginning and it is now more neglected than it was in 1928-1929.

III. Mental Testing in the Other Arab Countries

Apart from Iraq and Egypt, a mental test called,

" Ikhtibarat Akliya lil sufuf Al'Aliya Fil Madaris Al-Ibtida'yia"

اختبارات عقلية للصفوف العالية في المدارس الابتدائية

was adapted in Palestine for the higher classes in the elementary schools. The aim was to select students for secondary schools. It was made originally by Professor Godfrey Thompson and adapted to Palestinian environment by Jerome Farrell and Ahmad Sameh Al Khalidi in 1924 under the auspices of the Department of Education in Palestine. Nothing is available now about this study to indicate whether it was scientific or not. Therefore we cannot evaluate it.

In Syria no test has been attempted except one that was adapted by Dr. Kamel Nasry as a study for his Ph.D. degree, more than twenty-five years ago. Apart from that purpose it has not been used at all.

In Lebanon Professor George Shahla more than twenty years ago adapted "Terman Mental Ability Test" as a study for his M.A. degree. The language of the test was left in English for the purpose of offering objective test to select students coming from different countries as a supplement to college entrance examinations in the A.U.B.

IV. The Test Made in Sudan

The best adapted test in the Arabic language, which was carried on scientifically was in Sudan during the years 1946 to 1948 by G.C. Scott[1] under the auspices of the Gordon

---

(1) G.C. Scott, "Measuring Sudanese Intelligence", Psychological Abstracts, Vol. XXV, January, 1951, P. 32.
Also, The British Journal of Educational Psychology, Vol. XX, Part I, February 1950, P. 43.

Memorial College, Khartoum, for children of elementary and intermediate schools. Age 7-12 and 11-16.

Here is a brief presentation of the steps the author followed in the construction of his test in order to help the reader to compare this scientific study to the trials made in this field in the Arab World.

a) He started with an adaptation of the Al Qabbani Test[1] by changing the language from the colloquial Egyptian Arabic to colloquial Sudanese Arabic and to Sudanese environment. He administered it to 900 children of ages 10-16. The result was not satisfactory.[2] He discovered that the adaptation was not good;therefore he began to work again on it.

b) Thus he administered 400 items and only eighty-eight were selected. Moreover alterations in the wording and methods of presentation were made. Thus a new test came out - the "Second Oral Test". It was better than the first one[3] but he did not offer it for practical use.

c) He began in this step to work on the printed test. The first one was a selection and translation of six actual Moray House tests. As the first oral test it was done without experiment and the figures were even less satisfactory. Thus as a test it was useless.[4]

d) He started then to work on a second printed test which was a

---

(1) Ismail Al-Qabbani, "Ikhtibar Al-Zaka' Al-Ibtida'i", op. cit.

(2) G.C. Scott, "Measuring Sudanese Intelligence", op. cit. P. 43.

(3) Ibid., P. 44. (4) Ibid., P. 48.

printed version of the "Second Oral Test" (mentioned in Step 2). The author made it purposely this way in order to find out the effects of change from oral to visual presentation.(1) Consequently when he correlated the "Second Printed Test" with the "Second Oral Test" in grade V to VII he found that with improvement both in content and presentation a reliable printed test can be made for these grades. On the contrary in Grade IV he found an oral test was better for them.(2)

e) Hence he based on the second oral test two further tests the "Third Oral Test" for elementary schools and the "Third Printed Test" for intermediate schools. The coefficients of correlation of the "Third Oral Test" with the marks awarded in the intermediate school and with Entrance Examination were $r = \pm 0.82$, S.E. 0.057 and $r = \pm 0.61$ with S.E. 0.057. The reliability of the test was $r = \pm 0.94$ with S.E. 0.011. The author concluded that the test appeared to be sufficiently sound for use in Sudanese elementary schools.(3) Standardization was made on 780 children in grades III, IV and V in 14 schools.

As for the third Printed Test the realiability was $r = \pm 0.66$ S.E. 0.037 (for 236 boys in grade VIII). The reliability was fair but not so good $r = \pm 0.91$ S.E. 0.012 (for 171 children). Standardization was made on 689 pupils in grades V, VI, VII, VIII, and IX.(4)

---

(1) Ibid, p. 48.

(2) Ibid, PP. 48, 49.

(3) Ibid, P. 50.

(4) Ibid, P. 51.

There is also a new Printed test in Sudan, intended as an improvement on the Third Printed Test which is being worked on by Mr. Derar Saleh, Mr. Scott's colleague.

In the writer's opinion these tests tried in Sudan were the only tests in Arabic which have a sound scientific basis. Next comes "Ikhtibar Al-Zaka' Al-Ibtida'i" made by Mr. Al-Qabbani in Egypt. But all the rest have questionable value and cannot by any means be called standardized tests because the processes of adaptation, application and interpretation are not scientific and this is clear in the reports published about them.

Standardized tests become at present day an indispensable tool for scientific researches, therefore if the authorities in our country are seriously intending to solve the different problems confronting them in education, especially the adequacy of the actual educational system, there would be no way to dispense with the modern techniques of educational and psychological methods.

CHAPTER IV

CAN GROUP MENTAL TESTS PROVIDE A SOLUTION TO EXAMINATION PROBLEMS IN SYRIA?

I. From the study of the techniques used in the west for evaluating the students' abilities and capacities, and considering the meagre achievements in the Arab world in this field, we came to the conclusion that psychological tests are badly needed in this part of the world.

But to start entirely from the beginning is a very long process. Thus the best and shortest method, in our opinion, is to profit from the west's experiences in these matters by adapting the best tests available there.

Having the idea of adaptation, the writer started to study the tests found in the Department of Education in the A.U.B. in order to choose one, taking into consideration the following points:

a) The value of the test must be recognized in the west.

b) The administration and scoring of the test must be easy and possible within a short time. Thus a group test is more adequate to be chosen than an individual test.

c) Since the test represents a new type of examination for our students who are accustomed to the essay type of examination, it is advisable to choose a test in which the time limit for the whole test is not divisible for every sub-test apart.

Moreover this kind of test is more easily administered and the time more easily adapted than to go through the adaptation of the time limits of every sub test of the whole test.

d) Verbal tests are preferable to the non-verbal for our actual conditions, for three reasons:

1) Though the latter are sometimes said to be more free from the peculiarities of a special culture than the former, yet many adaptations are always necessary. Hence, their adaptation would be, in the writer's opinion, more difficult.

2) The use of this type of tests in the west itself, is still limited.(1)

3) Their content would be quite new for our students as a type of examination.

II. The Group Intelligence Tests Available in the A.U.B. in 1950 (Evaluation).

a) Revision of Army Alpha examination.(2)

The content of this test is generally adequate as other verbal group tests but the administration of this test is rather difficult. A limited time counted by seconds is designed for every subtest, e.g. the time allowed for the first sub test is 5 seconds. Moreover, the subjects have to move the papers upside down from time to time, which increases the students' anxiety and confusion.

---

(1) See Chapter II, P. 46.

(2) Elsie O. Bergman, Psychological Corporation, New York, 1935.

b) Culture-Free Test.[1] (80-90 minutes)

This is a non-verbal group intelligence test. It is designed to measure Spearman's "G" factor and to be relatively uninfluenced by verbal or cultural features. "Moreover, states L.S. Penrose, ability to do perceptual problems is said to be free from the influence of cultural and educational training. However, it is impossible to get away entirely from differences in performance due to familiarity or unfamiliarity of subjects with test situations in general."[2] David Wechsler says about this test, "The title of this test is misleading for, there is little evidence to show that the test is free from over all cultural influences, however defined."[3]

c) California Test of Mental Maturity.[4] (Two periods 45-50 minutes.)

(Many various short forms such as New California Short-Form Test of Mental Maturity, 1947, California Intelligence Test 1948).

Special features of these tests:

They yield a diagnostic profile showing mental age in separate tasks "language" and "non language" IQ's. Henry E. Garrett in criticizing these tests states "The IQ thought of as a developmental ratio, has little or no meaning when used with adults, and is always suspected when calculated from group tests.

---

(1) Raymond B. Cattell, The Psychological Corporation, N.Y. 1940.

(2) Oscar Krisen Buros, "Intelligence Tests", The Third Mental Measurements Year Book, Rutgers University Press, N.Y., 1949, P. 312: 228.

(3) Ibid, P. 313: 228.

(4) Elizabeth T. Sullivan, Willis W. Clark, & Ernest W. Tiegs, California Test Bureau, 1946.

It would have been far better if some other type of score had been employed. The score derived from the test battery as a whole is probably the most valid measure of mental level which the tests yield."(1)

Kuhlman on the other hand says, "The unabbreviated batteries are among the very best."(2)

According to our conditions the diagnostic profile cannot be used in the adapted test because it is in terms of Mental ages. Thus after adaptation process the test would loose one important feature of its characteristics. Moreover the large amount of verbal as well as non-verbal materials should be adapted the thing which adds another difficulty to its use. Besides, we quote Freeman in his conclusion about these tests, "It seems that these scales would find their best use in the hands of a clinical examiner rather than as instruments providing numerical data which have been established by the accepted standards now prevalent in psychological testing."(3)

d. Kuhlman-Anderson Intelligence Test(4) (40-60 minutes).

It is composed of variety of test materials, scaled over continuous range. It is relatively hard to administer. We quote Stanley S. Marsolf in this point, "Perhaps the biggest

---

(1) The Third Mental Measurements Year Book, op. cit. PP. 304 - 305 : 223.

(2) Cronbach, op. cit., P. 174.

(3) Freeman, Psychological Testing, op. cit., P. 216.

(4) F. Kuhlman & Rose G. Anderson, Educational Test Bureau, Minneapolis - Nashville, Philadelphia, 1947.

source of possible error in the use of these tests by the classroom teacher is the timing."[1]

e) Miller Mental Ability Test.[2]

No information whatsoever has been available to the writer about its value as it is looked upon by the Authorities in this field. Hence it has been eliminated.

f) Ohio State University Psychological Test.[3] (120 minutes, power test.

"It yields best predictions now available for an overall academic aptitude instrument at the college level." "The Ohio examination may be the result of a vicious circle which has over stressed verbal ability in education."[4]

The writer has found that it would take a long time in adaptation because it is composed of 150 items.

g) Otis Group Intelligence Scale[5]

It is for all literate persons including university students. It is difficult to adapt especially the "Narrative completion test". Moreover there is a time limit for every subtest, which thing is difficult to manage and adapt in our school.

There are still three other tests though they are good

---

(1) The Third Mental Measurements Year Book, op. cit., P.317: 236.

(2) W.S.Miller, World Book Co., Yonkers-on-Hudson, N.Y. 1922.

(3) Herbert, A. Toops, Ohio College Association Committee, 1947.

(4) Guilford, The Third Mental Measurements Year Book, op.cit,P.244.

(5) Arthur S. Otis, World Book Company, 1926.

yet some of their essential parts are missing, such as the manual. These tests are National Intelligence Test, Otis Quick Scoring Mental Ability Tests and Otis Self-administering Tests of Mental Ability.

III. Henmon-Nelson Test of Mental Ability - Form A[1]

This is the test which the writer has chosen for adaptation to Syrian students in the secondary school. Here is a summary of the reviewers comments about it leaving the detailed description of it to the coming pages.

Anne Anastasi declares, "Administration is very simple. The use of a single score based upon a hodgepodge of different types of items could be questioned, if one wanted a particularly fine or discriminative measure. But like other tests of its kind, this scale serves the practical purposes of

a) preliminary rapid exploration, and

b) rough classification of broad groups."[2]

August Dvorab comments, "The chief advantages of these tests lie in their ease of administration and scoring and that the relatively small number (90) of items require only 30 minutes of pupils' time."[3]

Howard Easley adds, "The content and standardization

---

(1) V.A.C. Henmon & M.J. Nelson, Houghton Mifflin Co., New York, 1931, the manual was revised in 1946.

(2) Oscar Krisen Buras, Intelligence Tests, the 1940 Mental MeasurementsYear Book, Rutgers University Press, New York, 1940, PP. 220, 221; 1393.

(3) Ibid, P. 222; 1393.

of these tests seem as satisfactory as those of the better group tests of intelligence but not strikingly more so."(1)

Mr. Guilford's evaluation of these tests is: "An examination of these group tests and the manuals and scoring keys that accompany them impresses one immediately with the kind of care and expertness with which one wishes all tests were constructed. The tests are appropriately entitled "Tests of Mental Ability" rather than "Tests of Intelligence" though the authors imply very clearly that they include a variety of mental operations, thus touching many areas of mental ability sampled in so-called tests of intelligence."(2)

In conclusion, we may say that this test is among the good tests according to the reviewers comments. Generally it fulfils to the four conditions(3) the writer mentioned before.

A. Description of Henmon-Nelson Intelligence Test - Form A, Grades 7-12.

This form parallels approximately the cycle of the secondary studies in the Syrian system. It consists of 90 items arranged in order of increasing difficulty. A wide variety of items is included such as:

- general information, such as, "Les Miserables" was written by:
  1. Hugo, 2. Faguet, 3. Moliere, 4. Martin, 5. Daudet;

---

(1) Ibid., P. 222 : 1393.

(2) Ibid., PP. 222, 223 : 1393.

(3) See Page 63 of this chapter.

- vocabulary includes items such as "Energetic" means about the same as, 1. sorry, 2. sleepy, 3. forceful, 4. happy, 5. brave;"
- logical selection such as, "A wheel always has 1. an axle, 2. circumference, 3. point, 4. a tire, 5. spokes;"
- arithmetical problems, "A silver shekel is worth about 60 ¢. How many shekels could one buy for $ 204.00? (1) 34, (2) 340, (3) 1284, (4) 3400, (5) 128."
- mixed sentences, "of term for is president a four the elected years. If these words were arranged to make a good sentence, what would be the next to the last work? 1. years, 2. elected, 3. president, 4. four, 5. term."
- verbal analogy, "Truth is to falshood as pride is to: 1. fear, 2. crime, 3. honor, 4. humility, 5. truth."
- geometrical analogy. It is based on the same principle as verbal analogy except that geometrical forms have replaced the words.
- classification, "Which word does not belong with the others? 1. Ida, 2. Paul, 3. Lucy, 4. Janet, 5. Edith."
- number completion, "1, 4, 7, 10, ..., ..., 19. What two numbers should be on the dotted lines? (1) 11 and 14, (2) 15 and 17, (3) 12 and 16, (4) 13 and 16, (5) 12 and 17."
- sentence completion, "It was raining too hard to ... out. A word for the blank is: 1. comment, 2. gather, 3. venture, 4. summon, 5. render."
- disarranged letters, "If the letters d g r e n a a r were arranged properly, they would spell: 1. renegade, 2. degenerate, 3. nagging, 4. arranged, 5. managed.

- proverb interpretation, "Better aim at a star than shoot down a well; you'll hit higher" means about the same as: 1. Hasty climbers have sudden falls. 2. Aim above the mark to hit the mark. 3. "Almost" kills no man. 4. Rash haste makes waste. 5. Might is not right."

Hence Spatial, as well as verbal and numerical materials, are employed. The different types of items are arranged in "omnibus" fashion, that is they are not grouped in sub tests each of one type, but are mixed up.

The authors state that originally two hundred and fifty items were prepared, and after they were submitted to teachers' criticisms only 202 items were retained, and they were divided into two forms and administered to 500 students. Only those items which differentiated between superior and inferior groups were selected. The justification for including items of general information was, as the authors state, their differentiation of students of high ability and students of low ability. Geometric analogies, included in this test proved to be among the most effective items for measuring students' abilities in U.S.A.

The validity of the original test:

The first criterion which has been used for validity was the comparison of the test with other tests which had proved effective as measures of mental ability such as the Otis Self Administering Test, Terman Group Test, the American Council Psychological Examination, Kuhlman Anderson Test and Illinois Intelligence Scale. The coefficients of correlation between

the Henmon Nelson test and each of these other tests ranged between .72 - .88.

In another study two forms of the Terman Test and two forms of the Henmon Nelson Test were administered to the same group of students and the IQ's averaged on the two forms. The correlation was .93 for one group of 144 pupils in one Junior High School and .94 for another group of 97 pupils in another Junior High School.

In a third study 179 pupils in the 8th grade were tested by means of the Terman group test. After four years these pupils who had reached the 12th grade, were tested by means of Henmon Nelson Test. The correlation between the IQ's in Terman Test and the percentile rank on the Henmon Nelson Test was .806.

Also correlations between means and standard deviations in the Otis and Terman, Kuhlman Anderson and Henmon Nelson Tests were obtained and they were found quite marked.

The reliability in the original test:

The reliability coefficient for each grade and for each age group were obtained. Each coefficient was based upon 100 cases selected at random. The method used was split half method. The coefficients were ranged between .882 and .936.

The norms:

As in every standardized test there are three tables representing norms based on scores made by about 30,000 pupils. One table is for determining IQ's, another for determining Mental ages and a third table showing the Percentile Rank of each score

for each grade for which the test is designed.

The total score in this test:

"It is influenced both by past school learning and ability to attack unfamiliar problems" as the authors state in the Manual. Speed is important and the total time is 30 minutes. The total score is equal to the sum of the right answers.

B. Adaptation of the Test

Hemmon-Nelson tests of Mental ability are constructed for students in U.S.A. Accordingly they are valid for these students in their special culture. Actually the criterion of validity of these tests, as it is the case with other intelligence tests, is the success in the social system the group live in. We quote Anastasi and Foley, "Intelligence tests measure certain abilities required for success in the particular culture in which they were developed."(1) Even the term success is also conditioned differently in different cultures. And since intelligence is only manifested through behavior, every culture encourages and stimulates certain kinds of behavior and neglects others.

Consequently the performance in a psychological test, of different groups living in different cultures, may be conditioned by different aspects of their life:

1) By the language or languages they speak;

2) By the traditions which develop certain types of beliefs,

---

(1) Anastasi and Foley, Differential Psychology, op. cit., P. 782.

ideals and standards which are inculcated since childhood;

3) By the educational system in the schools which stresses specific habits, skills and information;

4) By the specific type of training and parental care available to the children at home;

5) By the socio-economic level of houses and the facilities of intellectual advancement. Here nutrition, the diversity of objects in the home, and the variety of experiences to which the child is subjected, play an important role in his course of development;

6) By the emotional attitudes, interests, motivation, and preferences fostered by different surroundings. We quote Anastasi and Foley again, "It is apparent that many of the tests in current use cannot arouse the same emotional reaction in other cultures as they do in America."(1)

Hence the experiences of the students living in U.S.A. are different from those which Syrian students have. It is assumed therefore that the Herman-Nelson test of Mental ability will cancel its validity as an intelligence test for Syrian Arabic speaking pupils. Therefore both the text and the manual must be translated and adapted to the Arabic language and Syrian environment and afterwords a new validity for the test must be computed in the new conditions.

1 - The Nature of the Alterations in the Text

Modifications introduced in the content of the text

(1) Ibid, P. 734.

are due to two general causes:

a) the inadequacy of facts and information to the new environment,

b) the different syntax in the two languages.

Here are the types of those which come under the first cause:

1) Proverbs items which are strictly English or American in nature, e.g., "It is indeed an ill wind that blows no one good" means about the same as : Birds of a feather flock together; 2. Correspondence is half a presence; 3. Patience is the key of glory; 4. The calamities of one nation turn to the benefit of another; 5. The tongue is the neck's enemy.

2) Materials, names and measures which are not usually employed by the Syrian pupils, e.g., "How many feet are there in one rod and one and one-half yards? (1) 12, (2) 18, (3) 19, (4) 21, (5) 36."

"Asbestos is a product of : 1. trees; 2. petroleum; 3. mines; 4. wool; 5. cotton."

3) Information items which have social as well as educational value in U.S.A. For example, The Unites States entered the World War in : (1) 1914, (2) 1915, (3) 1916, (4) 1917, (5) 1918;

4) An item such as, "A book of the Old Testament is: 1. Matthew, 2. Luke, 3. Psalms, 4. Revelation, 5. Corinthians", was eliminated because Christianity is not the religion of the majority in Syria and thus it was replaced by one taken from arabic literature.

5) Most vocabulary items were changed for many reasons:

a) When they were translated into Arabic they have not kept their level of difficulty (they turned out to be much too easy, or much too difficult), e.g. "Superfluous is the opposite of: 1. certain, 2. essential, 3. ambiguous, 4. petite, 5. succinct."

b) When there is a certain trick either in the apparent construction or sound of the words chosen in order to tap the student's ability such as, "A recreant individual is: 1. young, 2. reborn, 3. smug, 4. happy, 5. cowardly." The item was eliminated and substituted by Arabic words based on the same principle.

c) When the vocabulary word has no synonyms in the Arabic language, in other words the writer could not find such synonyms for, e.g. "A minion is: 1. upright, 2. contrary, 3. servile, 4. poor, 5. impudent."

d) When some word in an analogy item did not stand for exactly the same meaning in the two languages, the analogy became wrong if it is not changed. Such as the word pride in the item "Truth is to falshood as pride is to: 1. fear, 2. crime, 3. honor, 4. humility, 5. truth."

The second cause of alteration: was due to the different forms of grammar and sentence construction in the two languages. This was mostly apparent in the disarranged sentences and disarranged letters, such as "of term for is president a four the elected years." If these words were arranged to make a good sentence,

what would be the next to the last word? 1. years, 2. elected, 3. president, 4. four, 5. term." And another item, "If the letters l e c r e a were arranged properly, they would spell: 1. reclaim, 2. earnest, 3. actual, 4. cereal, 5. leather." Sometimes the meaning was kept, and sometimes the item was changed totally.

The sum of the items which were adapted, was thirty-two, and all the rest of the items were translated verbally.

At last we may add that anyone who tries to produce or adapt a test for Arabic people, is confronted with two serious difficulties:

a) The difference between colloquial and classical Arabic especially as the latter does not represent exactly the mother tongue.
b) The actual Arabic dictionaries with their deficiency in giving the exact meaning of a word and its synonyms.

2. The Modifications in the Manual of Directions.

Attitudinal factors are serious in disturbing the performance of the students in a test and as it is very difficult to equate emotional reaction in different cultural groups, some modifications have been introduced in the manual so as to stimulate the Syrian students sufficiently and to get their participation.

The introduced alterations were found in three aspects:

a) In the number of exercises given to the students. Where in the original test there were three exercises representing

only three types of items, in the adapted one five more were added so as to familiarize the students with all the types of problems they would be asked to solve. Students in U.S.A. must have taken at least one test before reaching the tenth grade, therefore three exercises are sufficient for them; while Syrian students have not met such an experience before. Moreover, for an American school child the average intelligence test bears a close resemblance to his every day school work, he is consequently easily motivated to exert his best effort and to try to excel his fellows. Whereas this is not the case with our students in a Syrian high school who are always submitted to essay type examinations.

This addition of exercises is also justified on what is found in many representative tests where there is an exercise before every type of questions. Besides, the writer tried to get the best wordings (considering clearness and simplicity) in the explanations and directions.

b) The second type of alterations was in the content of directions. Inadequate presentation of the test not only reduces the score, but tends to invalidate the results, either because some students would not understand what they had to do, or because they would be uninterested, or afraid. Some "warming-up" sentences were introduced to get the best of students' participation, e.g., in the case of boys "every student should listen well and try his best so as to prove the high mental level of Syrian students and to raise the name of his school, because the

scores are going to be compared with that of the girls in the high school and with a group of American students."

c) Speed is another factor which is culturally conditioned at home, at school, and by the individual's surroundings as a whole. Therefore the writer got the idea that some alterations in the time limits would be needed. But every one who attempts this matter should ask himself what is the relationship between speed and intelligence? Group tests of general ability, aptitude, and educational achievement, with few exceptions, have time limits. Psychologists have for a long time questioned the correlation between speed and power.

The first important investigation into this matter was made with the Army Alpha Examination. This test was given twice to the same individuals, to the number of 510; once with the regulation time limits, and once with doubling this time. The coefficient of correlation between the two cases was found to be .965 which is extraordinarily high. In other studies the effect of time factor was entirely eliminated from the second testing, the coefficients of correlation were found to vary for different tests. Some r's were as low as .60, others were above .90. This means that speed has different weights in different scales. There are two divergent explanations in interpreting the high "r" between the two testings (single time and double time).

One group of psychologists finds that the high coefficient for a given test means that it is primarily a measure of power since the unlimited time in the second testing fails to

change the individuals' status. The other group states that the high "r" means that the test measures mostly speed and not power because the quick students come to a level which they cannot surpass even if they have a lot of time, while their capable slow friends profit from the additional time and consequently surpass them and the coefficient of correlation will be low.

Freeman found that any conclusion about such points depends upon the correlation between Speed and power. Because if the correlation, he says, between them is high, the increasing of time would give high "r" between the two testings whether the test was power or speed. The result of his experiment about this point was:

(1) Speed and power are correlated positively but the correlation is moderate and not high (.63).

(2) The correlation in a power test between the two testings (single time and double) is lower than in the case of speed test.(1)

Another experiment, stated by Ballard,(2) has recently been made with also Alpha examination by G.M. Ruch and Wilhelmine Koerth upon 122 freshmen. All those students were classified on the basis of four distinct intelligence examinations. The best students (52) and the lowest (70) were taken only while the average ones were eliminated. Three steps were followed in this study:

(1) Freeman, "Mental Tests, Their History, Principles and Applications," Houghton Mifflin Co., Boston, 1939, P. 265.

(2) Philip B. Ballard, "The New Examiner", University of London Press Ltd., London, 1948, PP. 104, 105.

a) The examinees were given the test under the time limits found in the manual. Lead pencils were used by the students.

b) Blue pencils were distributed and the subjects were allowed to do what corrections or additions they like under the same time limits in step 1.

c) After a few minutes rest, Red pencils were given out and the examinees were allowed to work on the tests until they felt that they had done their best. "r" between single and double time were found to be .966 and between single time and unlimited .945. We quote Ballard in interpreting the results, "All the results point to the conclusion that the prominance of the speed factor in the army tests does not invalidate the tests. There is abundant evidence to show that Speed and Power tend to go together."(1) On the other hand, Thorndike and Dr. Hunsicker got the conclusion that: "It seems unwise to attach much weight to speed, in intelligence examinations in general, except of course, in the case of tests where speed measures the speed of learning."(2) In most of the sounder scales today, the speed factor is minimized. However, many sources have shown that a curtailment of the time imposes no special disadvantages upon any general type of candidates though there are some exceptions, (the slow, cautious, timid or insecure) who are handicapped by

---

(1) Ibid., P. 106.

(2) E.L. Thorndike, E.O. Bergman, M.V. Cobb, Ella Woodyard, "The Measurements of Intelligence", Bureau of Publications, Teachers College, Columbia University, New York, PP. 400, 401.

imposition of time limits. Ballard states, "It can, at any rate, be shown that for a vast, and overwhelming majority, short period testing, when properly carried out, is as sound and as valid as long-period testing. There may be - there probably are - isolated instances where this is not true."(1) Then he concludes that the popular belief that slow readers are more profound and retain more in their readings is rejected by the actual psychological research. (Speed is not considered to be contrasted with accuracy).

As a compromise between these different opinions in Speed power studies, it is convenient for a test designer to put time limits for a test (due to the positive correlation between speed and power) but on the condition, as Thorndike and Cronbach state, that most of the given population can pass it within that time.(2)

Consequently the time given by the writer at first for the adapted test was 30 minutes, the regular time designed in the original test. But when it was found that only a few could finish the test within that time, it was extended 10 minutes more.(3)

C. The Sample Used in the Experiment

Since the original test has been constructed for American high school students from grade 7 - 12, the writer found that the fifth year in the Syrian secondary school may be taken

---

(1) Ballard, The New Examiner, op. cit., P. 107.

(2) Thorndike, The Measurement of Intelligence, op. cit., P. 405.

(3) Cronbach, Essentials of Psychological Testing, op. cit. P. 173.

for performing the test. This grade in the Syrian secondary school is usually reached after ten years of study and it is the year just after the Brevet class. And since every student cannot enter this fifth grade if he has not passed the Brevet examination, those students included in the sample were considered according to the prevailing system of education in Syria as having approximately the same scholastic background. Moreover there was an advantage for the writer in choosing this grade and this was: to be able to compare the scores of the students taking this test with their averages in the Brevet Examination, in addition to their teachers' marks during the actual academic year. As in every research work in psychology and education it is very difficult to perform the experiment on the whole population; therefore there must be a sample which consists of a limited number of individuals representing the population for which the test was put.

Many sampling techniques are usually employed in educational research and their use differs according to the different conditions. The techniques used in this study were those of the Simple Random Sampling method which are easy to control and follow and they are based on the laws of chance or probability. In other words the selection of the subjects according to the random sample is based on two things:

(a) every individual in the population has equal chance of being chosen;

(b) all selections are independent; that is the selection of one individual does not influence the choice of another.

A very small sample even if random, may not be representative of the population. Nevertheless it cannot be generalized as to the number of cases needed in a sample because that will depend upon the conditions of the particular problem.(1)

In our present study it was found that a random sample of one-hundred and forty-two students was suitable to yield approximately good results according to the time and facilities available to the writer. The procedure was carried on as follows:

A public secondary school for boys in Damascus was selected in a random way by taking the first school name which came first in the telephone book. But as there was only one public secondary school for girls which had a fifth grade it was necessary to take it. In this way we selected two secondary schools, one for boys and the other for girls.

The selection of boys and girls from the public schools without including private schools was justified on the grounds.

(1) That not all the headmasters of the private schools accepted the performing of the test on their students.

(2) The students in the private secondary schools, especially those who were in the fifth grade, were very few and of different backgrounds in their previous schooling. Therefore it was not advisable, in our opinion, to include them in that small sample.

Hence, having decided on the two schools and the grade there was

---

(1) For an extensive study of the sampling theory see Palmer O. Johnson, "Statistical Methods in Research", Prentice-Hall, Inc., New York, 1949, Chapter IX, also, Quinn McNemar, "Psychological Statistics", John Wiley and Sons, Inc., New York, 1948, Chapter V.

still another thing to be settled. In the girls' school there were three divisions of the fifth year, one scientific which consisted of (31) students, and two other art divisions one consisting of (43) students and the other of (41). The opposite thing was found in the boys' school where there was only one art division composed of (32) students and two science divisions consisting of (43) and (38) students.

Therefore the only science division found in the girls' school, and the only art division in the boys' school were taken. Then one of the two art divisions in the girls' school and one of the two science divisions in the boys' school were selected by choosing the division in which the second name of the first student on the register of the class, began with a letter which came first in the alphabetical order.

Thus in this study (142) students (72 girls and 70 boys) in the fifth year in the public secondary schools in Damascus were selected for the performing of the adapted test.

D. <u>The Administration of the Test.</u>

Two difficulties confronted us in that time:

(1) There was no single room in girls or boys high school which could be used to include all the girls or all the boys at the same time.

(2) It was impossible to give the test to the boys and girls at the same place because, it is against tradition in this country for boys and girls in the secondary cycle to meet at one hall.

Therefore the Education Council in the Syrian Ministry of Education helped in this study by providing a large hall in the Syrian University with confortable arm-chairs and tables for the administering of the test. It was a well-lighted room with two fire places (because the test was performed on February when it was very cold in Damascus.) This room was reserved for two days from eight to ten in the morning. On the first day the test was given to the girls and on the second day was given to the boys. However, the time of day is rarely an important factor and equally good results can be got at any hour if the individuals taken for the test are well motivated. But the writer preferred early in the morning before the students began their daily work, for them to be fresh and not tired, especially as those who were in the science divisions might have been more tired than the others in the art divisions during the day.

There were fourteen assistants and the writer, this means eight supervisers each day for 72 girls in the first day and 70 boys on the second day. One procedure was followed in the two testings. The directions and practice exercises of the test were given by the writer who had prepared for the testing by studying thoroughly the adapted directions designed to be put in the new manual.

A short talk was given to the assistants before the administering of the test and it consisted of the aim and the procedure which was to be used.

When the students were sitting calmly, the tables were

cleared of all materials and every student had a pencil, the writer began as follows:

1. "Today you have a new type of examination which contains short and interesting questions in eight pages. Every student should listen well and try his best so as to prove the high mental level of Syrian students, and to raise the name of his school. For the scores are going to be compared with those of the girls in the high school (the opposite thing is said to the girls) and with a group of American students."

2. "Do not write any work or number as an answer, since the answer is only to put a mark as I will explain to you later."

3. "For every question there is a limited time, therefore do not spend too much time on any one item."

4. "When you begin to write, no student is allowed totalk or to ask any questions. Any question of any kind is forbidden for I am going to explain to you the procedure in detail."

5. "Stop writing immediately when you are told to do so."

6. "To make the work easier and faster, fountain pens are not allowed, use only your pencils and when a pencil is broken during the examination do not speak but raise your hand and we will supply you with a good one."

7. "In a minute we are going to distribute the papers and they will be placed on your desk face down. Do not turn them over until I tell you to do so."

8. We distributed the papers.

9. Then the administrator of the test said, "Now turn your papers over so that you follow what I am reading." "We shall first fill out the blanks which are on the top of the first page (with a copy in hand, the writer indicated where these directions were found). After the word "pupil" write your name. (Pause). After the word "school" write the name of your school. (Pause). etc. " ... till every student filled out all the blanks about his class, division, the date, his age, etc.

10. Then the writer read the practice exercises slowly with a clear voice while the students were following her. At the end of each practice exercise, the writer paused while each student made the mark in the proper square.

11. At the end of the exercises the writer said, "Does everybody understand just how the test is to be done?" I answered only those questions which were reasonable and about the procedure which must be followed in doing the test. Then the writer said, "Is everybody ready?" Begin. I registered the time they began by hour, minutes and seconds on a sheet of paper before me, so as not to forget the exact time the students started the test.

The time in the original test was thirty minutes but when 27 minutes had passed only one girl had finished and when 30 minutes had passed two other girls had reached the end of the test. This means that three girls out of 72 had finished in 30 minutes. As we stated before ten minutes more were allowed

after we, the supervisors, had put a mark on the item that each girl had reached in 30 minutes. After 40 minutes I said, "Stop" and the papers were collected by the supervisors.

In the case of boys there was one who finished in 25 minutes and another in 28 minutes and three others accomplished it at the end of thirty minutes. Therefore the added ten minutes were needed here also. Thus as with girls so with the boys, the papers were marked after 30 minutes beside the last item attempted and at the end of the added ten minutes the papers were collected.

E. The Scoring of the Test.

The score of a student in this test is the total of the right answers and this is counted quickly in the original test with the help of the carbon paper which is underlying the squares marked by the students. Thus the tester has only to count the squares that are marked out on the inside pages.

In the adapted test, however, the examiner put a mark beside every right answer and then afterwards she counted the whole marked squares for every student. In this way the raw score for every student was obtained.

It should be noted, in this connection, that raw scores in psychology and education have not a constant value as the meter for example has in physical measurements. They have some limitations. Here are the most important of them:

1) Their meaning is relatively explained, e.g. a student who gets a score of 75 in a test cannot be considered as being

at the same mental level as another who gets 75 in another test, because the value of each score depends upon the scatter of other scores measured by the test.

2) The units in a mental test are not constant or uniform along the entire scale, e.g. the difference between a score of 30 and 31 is not equal to that between 80 and 81. Also it cannot be said that a student who gets a score of 60 is two times more intelligent than a student who gets 30. Even when two students have the same score in the test it cannot be said that they are equal.

3) There is no meaning to zero in these tests. Therefore when a student gets zero in an intelligence test this does not mean that he has no intelligence. Many conditions contribute to that.

Consequently no direct measurement and comparisons are possible in psychological testing. Therefore the indexes of relative rank become indispensable in every test so as to enable the examiner to evaluate the mental status of an individual. Unfortunately, all the norm tables found in the original test could not be used for a very vital reason: norms, as we have stated before, are averages taken from a group of individuals who are representative of the whole population for whom the test is intended. The performance of those representatives is dependent of course upon the conditions and opportunities available for them during their life in their American environment. Hence "unless the two populations who take the test, have been

subjected to similar cultural influences, for the same length of time, no comparison can be made."[1] This is of course impossible. This is why the different tables of norms, standardized in the United States for the original test, have no value in our study. Also to construct new tables of norms representing the Syrian students is very difficult. This demands a long time, and the cooperative efforts of a group of well-trained people in this field, who are helped financially by a rich agency.

Therefore the only way provided in our actual circumstances to judge the student's status is to rank him only approximately according to his group used as a sample for the test. This is the procedure which has been followed in this study as will be shown in the following chapter.

---

(1) S. Bieshenvel, "Psychological Tests", The Year Book of Education, Evans Brothers, Ltd., London, 1949, P. 104.

CHAPTER V

STATISTICAL METHODS AND EVALUATION OF THE RESULTS

The evaluation of the adapted test depends upon the results revealed by statistical techniques, these techniques are commonly known by those who work with psychological tests.

I. The first thing I began with was the analysis of the items upon which the validity and reliability of the test depend. I studied and analyzed the 90 items by using two criteria:

1) Most of the items should be of approximately 50 % level of difficulty with a few items scattered on both sides.
2) Every item must discriminate significantly between a superior and an inferior group.

Therefore four tables for the different divisions were prepared in which, the number of the students who failed in every item was found as the example shown in table I.

Then the percentage of failure for every division and then for all the divisions in every item were computed as it is summarized in table II.

The most important characteristic of an item is its discriminative powdr. And since every test as a whole must have this power, it is clear that this must be consistent from one item to another.

To study this characteristic in the actual test, the writer followed Kelley's procedure which consists of finding the

TABLE I - An example of the procedure followed in counting the number of students' failures for every item.

Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	...	Score
Student A	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	...	83
B	+	0	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	...	81
C	+	+	+	+	+	+	+	0	+	+	+	+	+	+	0	+	+	+	+	+	...	72
D	+	+	+	0	+	+	+	+	+	+	1	+	+	+	+	+	0	+	+	+	...	71
E	+	0	0	+	+	+	+	+	+	+	+	+	0	+	+	+	+	+	+	+	...	71
F	+	+	+	0	0	+	+	+	0	+	+	0	+	+	+	0	+	+	+	+	...	71
G	+	+	0	0	0	+	0	+	0	+	+	+	+	0	+	+	+	0	+	+	...	70
H	+	+	+	+	+	+	0	0	0	+	+	+	0	0	+	0	+	0	+	+	...	66
I	+	+	+	0	0	0	0	+	0	+	+	0	0	0	+	+	+	+	+	0	...	64
J																						
.																						
.																						
.																						
Number of failures	∴	2	2	4	3	1	3	2	4	∴	1	2	3	3	1	2	1	2	∴	1	...	

TABLE II - The Percentage of Failure for Every Item of the Test.

	Arts Girls	Science Girls	Arts Boys	Science Boys	Total	
1	10(41)	6(31)	11(32)	9(38)	36/142	25 %
2	13	10	10	9	42	29 %
3	8	6	9	4	27	19 %
4	5	2	2	2	11	7 %
5	8	4	5	4	21	14 %
6	3	2	6	3	14	9 %
7	9	3	12	2	26	18 %
8	16	17	17	14	64	45 %
9	4	2	7	3	16	11 %
10	14	8	11	4	37	26 %
11	5	1	4	3	13	9 %
12	25	17	9	13	64	45 %
13	4	3	6	11	24	16 %
14	10	7	15	4	36	25 %
15	30	22	24	28	104	73 %
16	6	0	5	6	17	11 %
17	8	5	20	15	48	33 %
18	27	18	18	10	73	51 %
19	6	2	8	2	18	12 %
20	16	7	14	9	46	32 %
21	0	0	0	2	2	1 %
22	2	7	5	3	17	12 %
23	1	1	2	1	5	3 %

TABLE III (Continued)

	Arts Girls	Science Girls	Arts Boys	Science Boys	Total	
24	3(41)	5(31)	3(32)	3(38)	14/142	9 %
25	1	1	0	2	4	3 %
26	23	15	17	20	75	52 %
27	7	1(30)	3	2	13/141	9 %
28	3	0	2	4	9	6 %
29	6	3	4	6	19	13 %
30	5	2	5	5	17	12 %
31	13	9	12	11	45	31 %
32	4	2	7	3	16	11 %
33	35	26	28	20	109	77 %
34	17	9	12	12	50	35 %
35	4	2	2	1	9	6 %
36	17	11	7	15	50	35 %
37	7	1	4	4	16	11 %
38	1	1	1	0	3	2 %
39	9	6	4	9	28	19 %
40	34	19	22	20	95	67 %
41	2	3	1	4	10	7 %
42	10	9	10	5	34	24 %
43	2	1	0	3	6	4 %
44	1	3	3	2	9	6 %
45	24	13	12	9	58	41 %
46	14	12	13	12	51	36 %
47	13	9	7	7(36)	36/139	25 %
48	14(40)	7	7	7	35/138	25 %
49	10(39)	9(29)	10	2	31/136	22 %

TABLE III (Continued)

	Arts Girls	Science Girls	Arts Boys	Science Boys	Total	
50	13(39)	10(29)	7(32)	19(36)	49/136	36 %
51	28	18	14	18	78/136	57 %
52	25	20	15(30)	12	72/134	53 %
53	5(37)	3	4	5	17/132	12 %
54	23(34)	20(28)	20	23	86/128	67 %
55	9	10	16	15	50/128	39 %
56	4(32)	7	5(29)	2	18/125	14 %
57	13	6(27)	4	11	34/124	27 %
58	26	19	18	19	82	66 %
59	3	2	5	4	14	11 %
60	19(31)	20(25)	17	15	71/121	58 %
61	13	14	18	17	62	51 %
62	13	11	6(28)	9	39/120	32 %
63	28 (30)	18(22)	24	30	100/116	86 %
64	19(29)	13	11(27)	20(35)	63/113	55 %
65	16	9	8(25)	10	43/111	38 %
66	2(26)	4	3	6	15/108	13 %
67	10(22)	15(21)	10	22(32)	57/100	57 %
68	10(18)	14	11	12	47/96	48 %
69	4	1(20)	7	8	20/95	21 %
70	1(17)	5(19)	6(24)	2	14/92	15 %
71	4(12)	6(18)	6(23)	5(31)	21/84	25 %
72	7(11)	8(16)	19	17	51/81	63 %
73	3	5(15)	12	16(30)	36/79	45 %
74	4(9)	3(14)	9	6	22/76	28 %
75	3	6	13	14(28)	36/74	48 %

TABLE III (Continued)

	Arts Girls	Science Girls	Arts Boys	Science Boys	Total	
76	3(8)	3(13)	4(23)	2(28)	12/72	16 %
77	5	6(12)	14	20	45/71	63 %
78	2	6	9(22)	13(27)	30/69	43 %
79	7	7	15	20	49	71 %
80	2	1	2	0	5	7 %
81	0	2(11)	3	0	5/68	7 %
82	3	5	9(21)	9	26/67	38 %
83	1	4	4	2(26)	11/66	16 %
84	1(7)	6	14	9	30/65	46 %
85	7	10	18	21(25)	56/64	87 %
86	0	0	2	1	3	4 %
87	0	0(9)	3	2	5/62	8 %
88	2	3	14(20)	11	30/61	49 %
89	6	5	16	13	40	65 %
90	4	3	8	11	26	42 %

highest 27 percent of individuals who had taken the test and the lowest 27 percent, on the basis of the test results.[1] Or according to our sample the highest 39 pupils and the lowest 39. Then the percentage of failures in each group for every item was computed and this is shown in table III.

According to this table, it was found that items(69), (82), (83), (85), (86) and (87) were missed by a smaller percentage of poor pupils than that of superior. Therefore they must be eliminated.

Items (54), (68) and (21) should be excluded also because the same percentage of inferior and superior groups failed to pass them.

But items (18), (23), (25), (28), (30), (35), (36), (38), (42), (43), (51), (61), (62), (63), (76), (78) have a slight discriminative power. They indicated that these items should be subjected to further critical study. Hence the papers of the two groups were reviewed, studied and compared with the original test. Here is the result:

1. Most of the rejected items were of vocabulary type (adapted or translated) and often they turned out to be very easy in the adapted test. Such as items, 82, 83, 86, 87, 21, 36, 38.
2. Information or facts type, such as items, (61), (18). The former in the original test was, "The United States entered the World War in: (1) 1914, (2) 1915, (3) 1916, (4) 1917, (5) 1918 ..." It became in the adapted test as follows,

---

(1) Freeman, Psychological Testing, op, cit., P. 26.

Item No.	The percentage of the superior group	The percentage of the inferior group	Item No.	The percentage of the superior group	The percentage of the inferior group
1	12	51	46	28	52
2	20	46	47	12	33
3	5	33	48	17	40
4	3	12	49	10	24
5	3	30	50	15	45
6	5	17	* 51	48	57
7	7	33	52	35	75
8	25	66	53	3	20
9	5	23	* 54	66	66
10	7	48	55	23	70
11	5	15	56	12	29
12	35	59	57	17	70
13	7	25	58	46	70
14	15	30	59	5	33
15	66	76	60	43	73
16	7	23	* 61	53	56
17	41	53	62	43	50
* 18	51	53	63	82	90
19	7	46	64	35	72
20	25	59	65	20	37
* 21	3	3	66	5	31
22	5	17	67	35	61
* 23	0	7	* 68	41	41
24	3	12	* 69	17	16
* 25	0	7	70	5	18
26	35	76	71	12	55
27	5	18	72	56	77
* 28	3	5	73	29	77
29	3	26	74	7	62
* 30	10	15	75	16	66
31	17	56	* 76	13	16
32	3	23	77	57	83
33	59	73	78	26	33
34	17	47	79	60	83
* 35	3	7	80	5	16
* 36	33	36	81	2	16
37	3	21	* 82	29	16
* 38	0	5	* 83	7	0
39	12	23	84	51	66
40	53	81	* 85	86	66
41	3	15	* 86	2	0
* 42	23	29	* 87	2	0
* 43	3	5	88	35	60
44	0	15	89	45	80
45	25	44	90	35	60

* Indicates the ineffective items.

TABLE III - Discriminative power of the adapted test items in percentage of failure by Superior and Inferior Groups.

"The Arab League was established in: (1) 1942, (2) 1943, (3) 1944, (4) 1945, (5) 1946 ... It appeared that even the most superior pupils could not locate the year. This may be because the Arab League is not sufficiently well known among Syrian students.

Most of the students (superior and inferior) in item (18) which was an original one, "water seeks its own ... a word for the blank is 1. money, 2. weight, 3. cold, 4. level, 5. length..." marked number (3). This indicated a lack of understanding of science which is a general weakness in our schools in Syria.

3. Geometrical analogies such as items (23), (25), (28) and (43) proved according to the experiment that they were very easy for the Syrian students. According to table II for example only 3% of the students failed to pass the first and the second one. Their discriminative power was also slight as indicated in table III. While the authors declare, "The items of this type included in the three forms of the test have proved to be among the most effective items of all."(1) We found them ineffective in our culture.

4. Item (62) which was translated from the following item, "Which word does not belong with the others? 1. rostrum, 2. structure, 3. cathedral, 4. edifice, 5. building", revealed that most of the students did not know the exact meaning of "cathedral" therefore they checked it as not

---

(1) V.A.C. Henmon and M.J. Nelson, Teachers' Manual, Houghton Mifflin Company, New York, 1946.

belonging to the series. This may be attributed partly to the Moslem environment. Therefore changing "cathedral" by "mosque" the item may become adequate.

5. In item (51) which was a translation of an original one, "A penitent person is always: 1. sorry, 2. carefree, 3. ill, 4. showy, 5. idiotic..." the personal opinion and culture play a great effect upon the response. In the manual the right answer was the word "sorry" while most of the students' responses were "carefree" and some few were "showy".

6. Item (63), "Circle is to ellipse as square is to: 1. oval, 2. cube, 3. curve, 4. circle, 5. diamond..." was changed into different content because the concept "ellipse" is not known by the Syrian students but in the highest classes in the secondary school. The adapted item proved to be very difficult (86 percent of the students failed in it) and its discriminative power was limited. Hence the adapted item may be replaced by another one which is very similar to the original one with some modification in wordings (names)

حسب التكييف الاول: "الهرم للمخروط كالمضلع الى: ١) المربع ٢) المكعب
٣) المنحني ٤) البيضوى ٥) الدائرة"

حسب الاقتراح الجديد: "الدائرة للشكل الاهليلجي كالمربع الى: ١) البيضوى
٢) المكعب ٣) المنحني ٤) الدائرة ٥) المعين"

Finally taking into consideration that the inadequate items, 76, 78, 82, 83, 85, 86, 87 were at the last part of the test (where only 6 out of 39 of the inferior group could arrive at while 38 out of 39 pupils of the superior group have accom-

plished the whole test) one may come to the conclusion that in these items errors played a greater effect among the inferior group than among the superior. For example item (86) was passed over by (1) out of 38 of the superior group and by (0) out of 6 of the inferior students. Thus the percentages were respectively 2 and 0. However the failure of 0 individuals out of 6 is not as a dependable result as the failure of 1 out of 38. Moreover the underestimated time of the test as will be found in the next chapter VI reflects another proof that the percentages of failures of the slow and inferior students at the end of the test would have been more accurate and representative if this time limit was sufficiently long. Shortly the inadequacy of the last items in the test is questionable.

Two choices were possible for the writer as a consequence of the analysis of the items:

1) Either to modify some of the inadequate items and to add a great number of new ones, then to subject them all to many successive corrections and try outs as it is followed in Standardized tests.
2) Or to eliminate the most ineffective items from the test if this elimination does not reduce too much the reliability. Since this latter choice is the easiest, the writer has begun with it. She eliminated the most inadequate items which were twenty, then she corrected the reliability coefficient after this elimination. It was reduced from 93 to 91 which still was high, higher than the original coefficient. Accordingly the writer preferred to follow this way though the first

method would have been better and more accurate. Neither time nor funds permitted the first alternative.

The last step in this analysis process was to arrange these remaining (70) items in a final form. Two things were considered in this step:

1) The level of difficulty of every item.

2) The omnibus type of arrangement of the original test.

Table IV represents an approximate arrangement that was tried for this purpose.

II. The reliability of the test[1]

The reliability in the original test was obtained by means of split-half (odd-even numbered items) method. The same technique was used in the adapted test. Besides it is easier and less time consuming than to go through the two other methods mentioned in Chapter II.[2] This is why it is universally applicable to all types of tests.

Table V shows the procedure followed in computing it.[3] The score of the odd items on the one hand and even items on the other were computed for every student. Then "r" the coefficient of correlation between these two sets of scores was

(1) For statistical computations we used: Henry E. Garrett, "Statistics in Psychology and Education", Longmans Green & Co., New York, 1935.
H. Sorenson, "Statistics for Students of Psychology and Education", McGraw Hill, New York, 1936.
R.L. Morton, "Laboratory Exercises in Educational Statistics", Silver, Burdett & Company, New York, 1928.

(2) See Chapter II, page 21. (3) In the appendix.

TABLE IV - The suggested arrangement of the remaining items.

41	44	81	80	27	4	9	6	11	24	59	16	53	22
19	66	15	56	29	30	5	37	32	7	70	39	14	3
48	49	74	1	10	47	2	57	62	71	20	31	34	17
46	65	50	39	82	45	90	73	12	75	8	84	78	88
26	52	60	64	67	40	58	77	72	89	79	15	33	63

calculated. Thus the two sets of scores were arranged in descending order. The highest score, in the column X representing the scores of odd-numbered items was given the rank 1 in the column named R_1, the third was given the rank 3, the fourth 4 and so on... But when two scores or more have equal rank each one of them was given the mean rank of these equal scores. The second set of scores which was put in the column Y represents the scores of even-numbered items and the rank of every score in this set was shown in the column named R_2. Then the difference between R_1 and R_2 was computed for every item and put in the column named (D). At last every difference was squared and the sum of all (D^2) was calculated and was equal to $\Sigma D^2 = 60411.5$. Putting this total in Spearman's rank difference: $P = 1 - \frac{6 \Sigma D^2}{N(N^2-1)}$

where n = number of individuals (142)

we got r = .874

As it was mentioned before, the whole test was longer and thus more reliable than either part of the full test. Therefore a statistical formula (Spearman-Brown) was used to correct the error: $${}_1r_2 = \frac{2\, r_{\frac{1}{2}\frac{1}{2}}}{1 + r_{\frac{1}{2}\frac{1}{2}}}$$

where $r_{\frac{1}{2}\frac{1}{2}}$ = is the correlation between the scores of the two halves, ${}_1r_2$ = is the reliability of the whole test. After the use of this formula the coefficient of reliability was raised from .874 to .932 which was a high reliability much higher than the reliability of the original test.

But since the result of the item analysis process was an elimination of 20 items, the writer was obliged to compute again the coefficient of reliability of the shortened test to see what would be the new reliability of the test. The formula used for this purpose was:

$$R_n = \frac{nr}{1 + (n-1)\,r}$$ where :

R_n = is the reliability of the shortened test.

n = means how many times the new test is shorter or (longer) than the original.

r = the original coefficient of reliability.

Hence in the actual test:

$$R_n = \frac{70/90\ (.93)}{1 + (7/9 - 1)\ .93} = .911$$

This last r (.911) was considered as the coefficient of reliability of the new adapted test.

III. <u>The validity of the test</u>.

I could not correlate my test with another known and valid tests because mine was the only one in existence. There remained before me only to correlate it with teachers' estimates of intelligence and with examination results, neither of which is conclusive. The trouble about teachers' estimates is that both their validity and their reliability are questionable.

As to their validity, most of the Syrian teachers in the secondary schools are not familiar with the technical meaning of intelligence and none has any experience of intelligence

testing. Some of the teachers include many abilities which are not related to intelligence. Accordingly many teachers for example do not differentiate between intelligence and school work but "they tend to estimate a child's intelligence according to the quality of his school work in the grade where he happens to be located."(1) Or they underestimate the intelligence of a shy student. Others do not take into account the age differences in the same class.

Briefly teachers' estimates cannot be relied upon especially in our schools and this is what Mr. Scott has discovered in his third printed Test in Sudan. We quote him, "I did not, this time, correlate the test scores with teachers' estimates, as I had become doubtful of their value".(2) The same thing happened to Ismail Al-Qabbani as is stated in the Manual of directions of his test "Ikhtibar Al Zaka' al-Ibtida'i".

Therefore I relied only upon the examination results. I correlated my test with two examinations. First with the Brevet Examination which is a State Examination. The test was given in December 1950 and the students had passed the Brevet examination either in June or in October. The other examination with which the test scores were correlated was the mid-year school examination which took place two months after testing.

Examination marks do not represent exactly the

---

(1) L. Terman, The Measurement of Intelligence, Houghton Mifflin Company, New York, 1916, pp. 73, 75, 76.

(2) The British Journal of Educational Psychology, op. cit., p. 51.

intelligence of a student because many different factors enter into them, such as lack of time or interest, or perhaps poor health, teacher proficiency, the student's personality or manner, etc..

Kuhlmann and Anderson state that "validity coefficients are misleading since a suitable criterion is not available. School marks cannot be used because they are contaminated by other factors".(1) But Stanley S. Marzolf reports, "while it is true that school marks do not constitute a perfectly desirable criterion, we have a right to expect an intelligence test to show some relation to them. Certainly, correlations of the order of .60 have been found with sufficient frequency to cause us to view with suspicion any test with a validity coefficient much below this level".(2)

Indeed psychologists have not obtained the same degree of correlation in their studies of intelligence tests and school marks. This variation may be due, as some state, to many factors, such as, the homogeneity or heterogeneity of the group, the reliability of the educational rating, the range of ages included in any one group, the thoroughness with which education and intelligence have been measured in each case, and other similar factors. "In general however, the results clearly indicate that the thing measured by our intelligence test is one of the factors making for success in school work".(3)

---

(1) The Third Mental Measurements Year Book. op. cit., p. 318.

(2) Ibid., p. 318.

(3) R. Pintner, "Intelligence Testing Methods and Results", Henry Holt and Company, New York 1923, p. 257.

Thus - the correlation of Binet IQ and school achievement has been found to vary from about .40 to .75,[1] for Terman in constructing the Stanford-Binet scale, used school success as one check upon the validity of these tests.[2]

Pintner has compiled the correlations reported by fourteen authors between general intelligence test scores and school marks gained by high-school pupils. These r's which were based upon varying numbers of students, the largest group being 5,748 high school seniors, vary from .28 to .60 the most frequent between .30-.60, the mean being .46.[3]

As to the validity of the original Henmon-Nelson Tests of Mental Ability, it was checked against scholastic achievement as well as against scores on other common intelligence tests. In some of the college groups, correlations were reported between test scores and grades in various courses, as well as composite first term grades. The latter correlation was .60, the others ranging from .46 to .60. Quoting Anastasi, "These correlations compare fairly well with these obtained between most scholastic aptitude tests and college grades".[4] Going back to the adapted Henmon-Nelson Test, every part of the four divisions: girls (Arts), girls (Science)

---

(1) H. Garrett and M.R. Schneck, Psychological Tests, Methods, and Results, Harper & Brothers, New York, 1933, p. 27.

(2) Terman, the Measurement of Intelligence, Houghton Mifflin Company, New York, 1916, p. 74.

(3) Pintner, op. cit., p. 256.

(4) The 1940 Mental MeasurementsYear Book, Op. cit., pp. 220-221: 1937.

boys (Arts) and boys (Science) was treated apart in correlating test scores with mid-year school examination, and Brevet examination, because the teachers were different in the four divisions and the subject matter varied from an art to science division.

Few subject matters were selected for the averages representing the teachers' marks in the mid-year school examination on the basis that they must be taught three times a week so as not to go in the long marks weighting process. For example Arabic literature, English, History and Physics were chosen for Arts divisions, and Arabic literature, English, Mathematics (Algebra, Geometry, Trigonometry) Physics and Chemistry were taken for science divisions. Composition and Logic were not included in the average because they were taught twice a week, though composition marks correlated high with the Binet test, according to a study made by Burt in the high school for a large number of children age 7-14 r = .63.(1) But in the case of the Brevet examination the average of all the subject-matters was taken without taking into consideration the hours spent on every subject which was a mistake according to the writer's opinion. We have to notice however, that in the Brevet grade there is no such distinction as science or art divisions but most of the teachers were different in the various divisions.

The coefficients of correlation of the adapted test with Brevet examination and School Mid-Year examination ranged

---

(1) Pintner, op. cit., pp. 256

from .04 to .54 as it is shown in diagrams (1), (2), (3), (4), (5), (6), (7), (8)[1] and the results were summarized in Table V.

Three out of eight correlations were inadequate and very low (.04, .08, .12) according to what has been stated in the above mentioned discussion about this point. Moreover the coefficients were generally lower with the Brevet examination than with teachers' marks and lower with girls than with boys. The latter result may be the effect of the difference in teachers between boys and girls schools. Besides, girls were much more afraid and emotionally disturbed than boys when the test was administered. But as for the first generalization (correlations with the Brevet averages being lower than with teachers' marks), it may be the outcome of either not taking into consideration the hours spent for every subject in the Brevet averages or the Brevet examination was less accurate and reliable in judging students' ability than the teachers' marks in the Mid-year School examination, or the result of both causes together.

Hence I correlated the marks of the Mid-Year School examination with the Brevet grades of the four divisions in order to see to what extent these two examinations coincide with themselves in judging students' ability, which thing may clarify the correlations found with the test. The results were, as it is indicated in Table V, .34, .50, .52, and .58.[2] This points

---

(1) See the Appendix.

(2) See also for more details, diagrams (9), (10), (11) and (12) in the appendix.

	Girls (Arts)	Girls (Science)	Boys (Arts)	Boys (Science)
Test and Teachers' Marks	.12	.36	.54	.26
Test & Brevet Examination	.08	.04	.46	.50
Teachers' Marks & The Brevet	.52	.34	.58	.50

TABLE V - The coefficients of validity of the adapted test.

out that the correlations between the two examinations are moderate and not as high as it should be because the aim and the nature of the questions in these examinations are very much alike. However, these correlations are higher than those found between the test scores and the Brevet Examination on one hand and test scores and Mid-Year School Examination on another hand. This result was expected because most of the teachers who put the marks in school examinations were usually called to correct the Brevet Examination papers. On the other hand the types and principles used in setting the Brevet Examination were usually followed in School Examinations which ordinarily prepare students for State Examinations.

IV. The Frequency Distribution (1)

Since test scores cannot be comprehended and interpreted but relatively, it is necessary to locate every pupil's standing in the group which has taken the test. This demands the use of statistical methods which summarize and organize the quantitative facts in order to facilitate their treatment and to reveal significant trends of the group's performance.

Therefore the first step taken, was to arrange the scores of the test in a grouped frequency distribution in which the scores of (142) students ranging from 17 to 83 were

---

(1) Helen M. Walker, "Elementary Statistical Methods", Henry Holt and Company, New York, 1948, Chapter III.
Also Harold Ordway Rugg, "Statistical Methods applied to Education", Houghton Company, Boston, 1917.

grouped into class intervals of five points.

In making such a distribution four steps were followed:

a) The determination of the range:

It is the difference between the highest score and the lowest and it was equal to 66.

b) The selection of the class interval:

It is the size of the groups into which the scores are to be classified. To do this, the range was divided first by (10) intervals which gives the largest class interval that can be used and then the range was divided by (20) which gives the smallest size of interval which can be used. Thus the size of the interval must be somewhere between 4 and 7. This means that the size of the interval might be 4, 5, 6, 7, but since (5) is a very convenient number to handle, it was chosen. Accordingly, the number of the intervals was 15.

c) Determining the integral limits of the intervals:

As the number of units in the interval was an odd number (5), to determine the integral limits of each interval the multiple of five nearest to the highest score was taken and found to be 85. This figure was considered consequently as the middle score in the highest interval such as, 83 84 85 86 87 and thus the integral limits of that interval must be 83-87. The limits of the other intervals followed automatically. Then these intervals were written in a descending

order in the first column headed (S) in table VI.

d) Making the tabulation:

In the second column headed (T) a short vertical line was drawn, for each score, in a line with the class where it fell. This made it easier to count the frequency in each class. The Frequency column abbreviated f, gives the number of scores in each interval. The third column (Cum. f) represents the cumulative frequencies which were obtained by adding the scores serially beginning with that on interval 13-17. (Table VI explains this clearly).

The facts brought out by this frequency distribution can be made more vivid if presented pictorially by means of a frequency graph. Hence the data of table VI were shown in graphic form in diagram (13). The horizontal axis represents the scores or in other words the intervals; the vertical axis shows the frequency or number of cases falling within each interval.

S Intervals	T	f	Cf	Steps in making the table
83 - 87	\	1	142	
78 - 82	|	1	141	
73 - 77	𝍸	5	140	
68 - 72	𝍸 𝍸 ||	12	135	
63 - 67	𝍸 𝍸 |	11	123	
58 - 62	𝍸 𝍸 𝍸 |	16	112	Step 1. Determining the range Highest score 83 Lowest score 17 Diff. 66
53 - 57	𝍸 𝍸 |||	13	96	
48 - 52	𝍸 𝍸 𝍸 𝍸	20	83	Step 2. Selecting the class interval 66÷10=6.6 66÷20=3.3 (5 is chosen)
43 - 47	𝍸 𝍸 𝍸 𝍸 𝍸	25	63	
38 - 42	𝍸 𝍸 𝍸	15	38	
33 - 37	𝍸 𝍸 |||	13	23	Step 3. 5x17 = 85 83 84 85 86 87
28 - 32	||||	4	10	
23 - 27	|||	3	6	
18 - 22	||	2	3	
13 - 17	|	1	1	
		N=142		

Table VI. A Grouped Frequency Distribution for the adapted test results.

The dotted curve represents the smoothing of the frequency polygon. It is based on Table VII, in which the average of every three adjacent frequencies, found in the column headed f, was taken by multiplying the middle one by 2 and dividing the sum by 4. The results were the new frequencies found in the third column of the table VII.

f	The procedure	The Expected frequency
1	2 + 1 = 3 ÷ 4 =	.75
1	2 + 1 + 5 = 8 ÷ 4 =	2
5	10 + 1 + 12 = 23 ÷ 4 =	5.75
12	24 + 5 + 11 = 40 ÷ 4 =	10
11	22 + 12 + 16 = 50 ÷ 4 =	12.5
16	32 + 11 + 13 = 56 ÷ 4 =	14
13	26 + 16 + 20 = 62 ÷ 4 =	15.5
20	40 + 13 + 25 = 78 ÷ 4 =	19.5
25	50 + 20 + 15 = 85 ÷ 4 =	21.25
15	30 + 25 + 13 = 68 ÷ 4 =	17
13	26 + 15 + 4 = 45 ÷ 4 =	11.25
4	8 + 13 + 3 = 24 ÷ 4 =	6
3	6 + 4 + 2 = 12 ÷ 4 =	3
2	4 + 3 + 1 = 8 ÷ 4 =	2
1	2 + 2 = 4 ÷ 4 =	1

TABLE VII - The smoothing process of the Frequency Polygon.

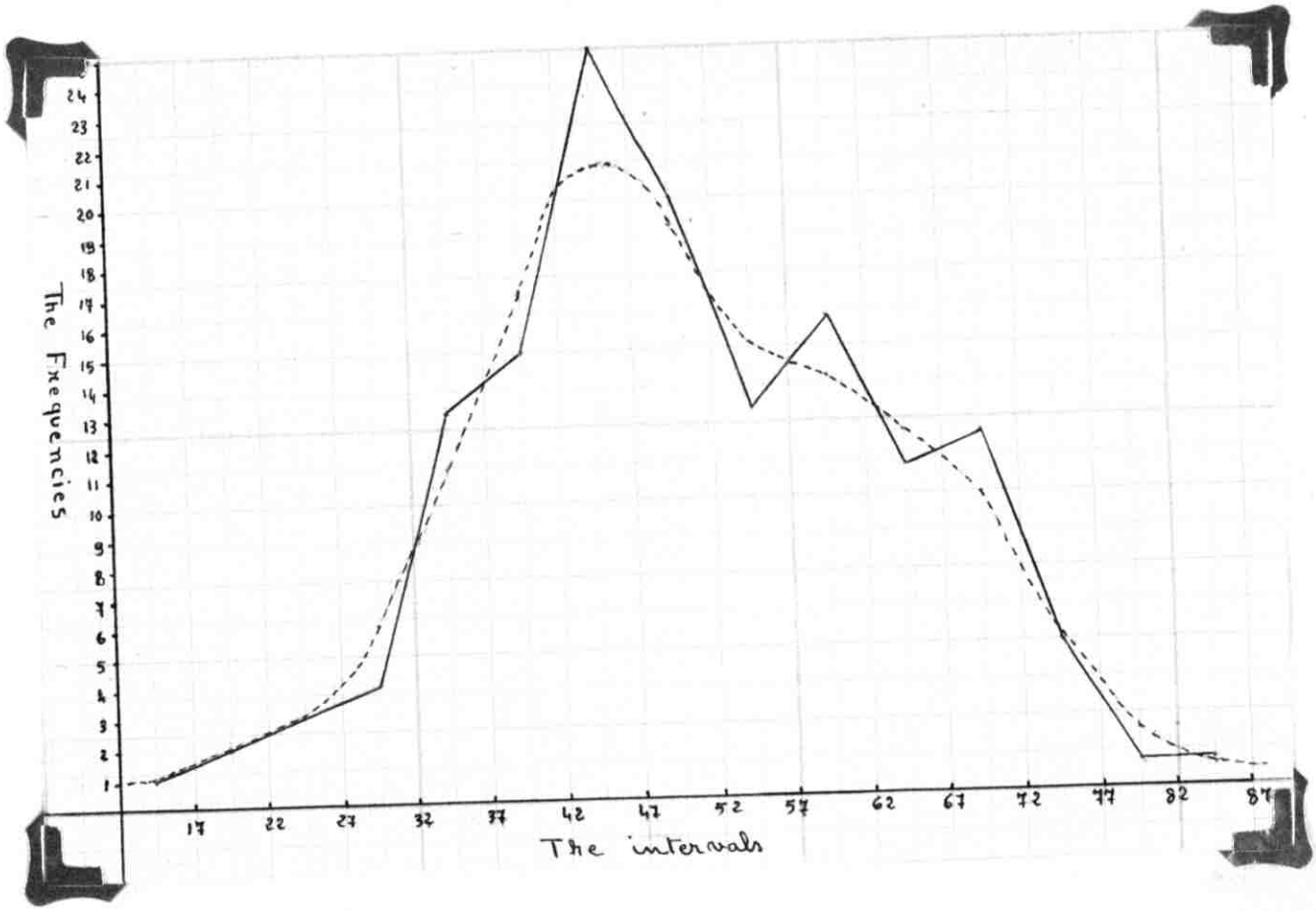


Diagram 13. A Frequency Polygon representing the distribution of the scores taken from the adapted test in 142 Senior high school students.

Another form for representing the above mentioned Frequency Distribution is the Cumulative Frequency Curve in which the cumulative frequencies found in Table VI were plotted against each interval instead of the frequencies. This is shown in Diagram 14.

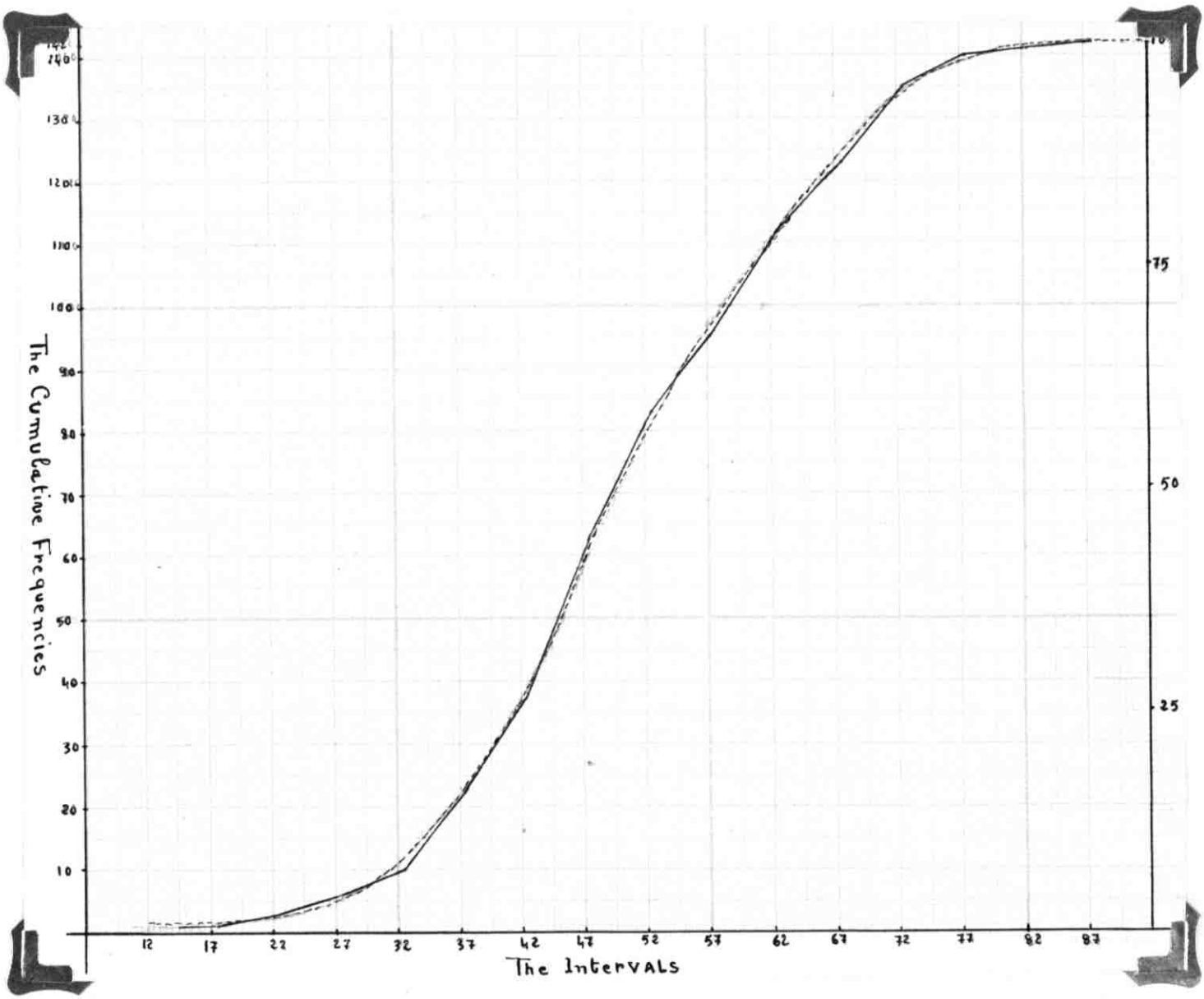


Diagram 14. The Ogive or Cumulative Frequency Curve representing the scores distribution of the adapted test. (See Table VI).

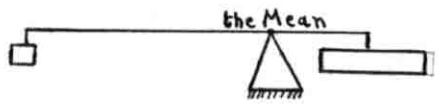
V. Measures of Average or Central Tendency.

A characteristic of most Frequency Tables is the tendency for the scores to concentrate somewhere near the center. Therefore the first and most important measure in statistical analysis is the location of the concentration point on the scale, that point which best represents the whole distribution.

Statisticians employ three common averages. These are the arithmetic mean, the median and the mode.

a) The Arithmetic Mean

We can compute the mean merely by obtaining the sum of the measures and dividing it by their number. The measure so obtained is then the value that each individual would have if all shared equally. Expressed in physical terms, it may be thought of as the point at which the fulcrum must be placed in order to balance the scale.



The steps the writer followed in computing the mean were as follows:

Step 1. A mean was assumed at the mid point of interval 48-52; 50 was assumed to be the mean M' (any interval might be taken).

Step 2. The deviations from the assumed mean were laid off. The plus deviations (see Table VIII) indicate how many intervals frequencies were above the assumed mean, and

Intervals	f	x	fx	fx^2
83 - 87	1	7	7	49
78 - 82	1	6	6	36
73 - 77	5	5	25	125
68 - 72	12	4	48	192
63 - 67	11	3	33	99
58 - 62	16	2	32	64
53 - 57	13	1	13	13
			+164	
48 - 52	20	0	0	0
43 - 47	25	-1	-25	25
38 - 42	15	-2	-30	60
33 - 37	13	-3	-39	117
28 - 32	4	-4	-16	64
23 - 27	3	-5	-15	75
18 - 22	2	-6	-12	72
13 - 17	1	-7	-7	49
			-144	
	N=142		Σfx=20	Σfx^2=1040

Table VIII. Illustrates some of the calculation of the mean M, Median and Standard deviation.

minus deviations indicate how many classes frequencies were below the assumed mean. This is shown in the column headed x.

Step 3. Each f (frequency) was multiplied by its corresponding x (deviation).

This is found in the column headed fx.

Step 4. The algebraic sum of the fx column (Σfx) was obtained.

Since the plus value (+164) exceeded the minus values (-144), the sum was plus 20.

Step 5. The correction for the assumed mean was determined.

Thus the sum of the fx column was divided by N, and multiplied by the size of the interval (i) which was 5 units.

20/142 = .14, the correction in terms of intervals.

.14 x 5 = .70, the correction in terms of scores units.

Step 6. The true mean was obtained by adding the correction to the assumed mean.

Here the assumed mean is 50 and the correction is .70. Thus the true mean is 50.70.

Briefly the formula for computing the mean from an assumed mean in a frequency table is:

$$M = M' + \frac{(\Sigma fx)\, i}{N}$$

$$M = 50 + \frac{20 \times 5}{142} = 50.70$$

b) The Median

It is the mid point in the distribution or that point

below or above which half of the number of scores lie. It is perhaps the most widely used average in educational measurement. The steps followed in obtaining it may be summarized in the following:

Step 1. The N/2 was obtained by dividing the total number of frequencies or cases by 2. $\frac{N}{2} = \frac{142}{2} = 71.$

Step 2. An approximate median was located: Beginning at the low end of the distribution, the frequencies in the frequency (f) column in Table VIII were counted up as far as possible without passing N/2 = 71. When the interval 43-47 was reached the frequencies became 63 which were lower than 71. Therefore the median must be in the next interval, the 48-52 and the real lower limit 47.5 of this interval was taken as the assumed median.

Step 3. The correction needed was determined: The total 63 obtained in step 2 was substracted from N/2 or 71-63 = 8. This shows that eight scores more or units were needed to obtain the required half. These scores must come out of the next interval 48-52 where there was a frequency of 20. That is we have to go 8/20 of the distance into the next interval. As the size of the interval was 5 units, this means 8/20 of 5, or 2 which was the correction needed.

Step 4. The true median then was obtained by adding the correction to the approximate median. In this case [illegible] 47.5 + 2 = 49.5, the median.

Putting the above process in a formula, we obtain:

$$Mdn = \ell + \frac{(N/2 - Fe)\ 1}{f\ p} \qquad \text{or}$$

$$47.5 + \frac{(71 - 63)\ 5}{20} = 49.5$$

By the help of the median every pupil in the distribution can be located as in the higher half above the median or in the lower half below the median.

Still the position of any pupil in this distribution may be more accurately described by indicating the percentage of pupils who fall below him. The points that divide a distribution into 100 equal divisions, are called percentiles. The percentile rank for every pupil can be read in diagram (14) which represents an Ogive in which the vertical line at the right side was divided into 4 equal parts (quartiles) that help in reading the scores in terms of percentiles from the curve and vice versa.

c) <u>The Mode</u>

It is the commonest score in a group or that point in the horizontal axis where the frequency is greatest. The mode is often regarded as the poorest average and it is computed from the formula:

$$Mo = 3Mdn - 2M$$

where Mo = the Mode.

Mdn = the median

M = the arithmetic mean

In our study:

$$Mo = 3 \times 49.5 - 2 \times 50.70$$

or $Mo = \underline{47.1}$

Finding the value of arithmetic mean (50.7), the median (49.5) and the mode (47.1) near each other, one can guess that the curve representing this distribution is approximately normal because in an ideal normal probability curve these three points coincide. This is what was found in diagram (13).

VI. Measures of Variability or Scatter.

The second problem in the statistical analysis of our test data was to determine the variability of the scores because central tendency measures do not describe completely the distribution. They do not show the extent to which the scores tend to spread above or below the average point. The Standard Deviation is the most important measure of the variability of test scores. "It depends on the test material and may vary from 13 up to perhaps 25".(1) A small Standard deciation means that the group has a small variability or it is relatively homogenous. It is generally defined as the square root of the mean of the squares of the deviations of the scores from their mean. Or briefly it is represented in the formula S.D. = $\sqrt{\Sigma x^2/N}$ where x = the deviation of each score from the mean.

It may sometimes be defined as that distance above and below the mean that in a normal distribution curve includes 68.26 percent of the scores, or approximately two thirds of the total scores.

---

(1) The Third Mental Measurements Year Book, op. cit., P. 316.

The formula for the Standard deviation from which it was computed in our study is:

$$= i\sqrt{\frac{\Sigma fx^2}{N} - c^2} \quad \text{where}$$

C = the correction of the guessed mean in terms of the interval as a unit

Thus the formula may be :

$$= i\sqrt{\frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx^2}{N}\right)} \quad \text{where}$$

x = is the deviation of an interval from the interval in which the guessed mean is considered.

Hence in our actual test according to table VIII:

$$= 5\sqrt{\frac{1040}{142} - \left(\frac{20}{142}\right)^2} \quad \text{or}$$

$$= 5\sqrt{7.3239 - .0198} \quad \text{or}$$

$$= 5\sqrt{7.3041} \quad \text{or} \quad = 13.51$$ which means according to the Third Mental Measurements Year Book[1] that our tested group is approximately a homogeneous group.

To test the extent of normality in the actual distribution the mean deviation must be computed. It is the mean deviation of the various scores from the mean. The formula of the mean deviation for a grouped frequency distribution is:

$$MD = \frac{i\,(\Sigma fx)}{N} + \frac{c^1\,(fa - fb)}{N} \quad \text{where}$$

x = the deviation of every interval from the mid-point of the interval containing the mean.

c^1 = is the result of subtracting the mean from the mid-point of its interval.

---

(1) Ibid, P. 316.

fa = the total of the frequencies of the intervals whose mid-points are above the mean.

fb = is the sum of the frequencies of the intervals whose mid-points are below the mean.

In our study according to Table VIII:

164 + 144 = 308 the sum of fx (treating them all positively)

1 + 1 + 5 + 12 + 11 + 16 + 13 = 59 = fa

20 + 25 + 15 + 13 + 4 + 3 + 2 + 1 = 83 = fb

50 - 50.70 = -.7 = c^1

Thus $MD = \frac{5 \times 308}{142} + \frac{-.7 \times -24}{142}$

or MD = 10.8 + .11 = 10.91

Hence MD/SD = 10.91/13.51 = 0.807 or 80.7 % our curve is near to normality. Since in a normal distribution the M.D is about 8/10 of one S.D.(1) This supports the same result found when the mode, mean and median were computed.

VII. The Sampling Error of the Measures of the Adapted Test.(2)

Statistical formulas are based upon certain assumptions which may be not fully met in our actual study. Since the test was performed on a limited number of individuals for whom it was constructed. Therefore the measures of central tendency and variability cannot be considered as representing exactly those of the whole population. Some errors or uncontrolled irrelevant factors always enter into the test results,

---

(1) Peters & VanWoorhis, "Statistical Procedures and their Mathematical Bases", McGraw Hill Book Co., New York, 1940, P.81.

(2) Helen M. Walker, op. cit., Chapter XV.

from the measuring instrument itself, from the subjects or from errors of sampling.

Therefore it has been shown by experience that when a large number of samples of the same size are taken, by the same random way and drawn from the same population for whom the test is intended, then their means are plotted, these means form a normal curve and the mean of these means will be approximately near the true mean of the whole population. The Standard Deviation of these means is then called the Standard Error. Therefore when the Standard Deviation (or the Standard Error) of these means is small every mean of these samples can be considered as an approximation of the true mean. But if the Standard Error of these means is large no one can be taken.

Consequently many formulas have been put to help every investigator to compute the Standard Errors of different measures of his study. Here are some of them applied to our study:

$$\delta_M = \frac{\delta}{\sqrt{N}} = \frac{13.51}{\sqrt{142}} = 1.13$$ The Standard Error of our mean.

$$\delta_{Mdn} = 5/4\ \delta_M = 5/4 \times 1.13 = 1.41$$ is the S.E of our Median.

$$\delta_\delta = \frac{\delta}{\sqrt{2N}} = .707\ \delta_M$$

or $\delta_\delta = .707 \times 1.13 = .79$ is the S.E. of our Standard Deviation.

Taking the first formula of this group as an example for explaining them we can say that the S.E. (1.13) means that in two cases out of three the true mean will lie between 50.70 $\pm$ 1.13 and in 99 cases out of 100 the true mean will be between 50.70 $\pm$ 3 x 1.13 or between 47.31 and 54.09 and so on for the other measures.

CHAPTER VI

THE EVALUATION OF THE ADAPTED TEST

After reviewing this study one comes to the conclusion that this adapted test in its actual content although representing a step in the right direction, is in need of further study and refinement. In other words it is still in its preliminary stage of try-out process and without further experiments it cannot be put in the hands of teachers for use.

Thus as a test it is still of questionable value. However many of the limitations of the actual test may have come out in the relatively short time which has been available to the writer for this study (one academic year). And it is well known that a study like this to be valid, needs years and years of hard work. Therefore taking this test as a seed or a first step of a test for Syrian secondary schools many improvements can be achieved in future and this study would be valuable and indispensable for that purpose.

Some limitations are found in the actual test: 1) in the sample, 2) in the items, 3) in the time limits, 4) and in the validity.

I. <u>The sample</u>:

The sample used for this test was too small when compared to that of a standardized test. It is composed only of 142 students, and even for a study like this Rugg states that, "500 cases are necessary to give a thoroughly continuous distribution

(smooth)."[1] But to take an adequate sample even if it was around 500 students and to do all the necessary statistical computation would have taken the writer a very long time especially as there is no computing machine in the Education Department in the A.U.B. for this purpose.

Moreover the selection of our sample, though it was by a good random procedure, yet it was in the writer's opinion somewhat biased. It is true that the selection of the boys' school in Damascus was at random, and this school receives students from different districts in Damascus and from various cities in Syria; also no selection is made in accepting the Post Brevet Students in this school but those who come first have priority over others. Yet, only two central schools in Damascus are not enough for a more developed test. Therefore we cannot say at all that our sample is representative of the 5th grade students in the Syrian Secondary Schools.

Another question still confronts us and needs clarification in evaluating the sample of the test and this is: What is the effect of the student's age in this cycle (Senior High School) upon the testing results ? Should the writer have to take into consideration the student's age in choosing the sample ? Here we come to a big problem still considered a debatable question and this is the age of adult intelligence. The range of age in our study was between 15-20 years and here some questions would arise: On what bases can we compare our

---

(1) Harold Rugg - A Primer of Graphics and Statistics for Teachers, Houghton Mifflin Company, Cambridge 1925, P. 62

students falling in a range of ages 15-20 ? Can we say in that case, for example, that these students who are 15 years old and in the middle of the distribution are less intelligent than those who are 18 and in the highest end of the distribution?

Two problems connected with this point have been studied by psychologists: The Mental Growth and the Consistency of the IQ.

A. As for the first one, psychologists as well as physiologists have tried to plot a curve for mental growth in order to find how mental growth develops, when it stops and how rapidly it decreases ? There is no general agreement among psychologists upon the maximum point the adults reach in mental growth. Some psychologists put that point as low as 14 while others say that it increases up to age 22. This is due partly to the nature of tests which do not give valid measures of intelligence at various age levels and partly to the nature of intelligence which is not yet definite. Freeman says that the cause of these low adult levels has two reasons:

1) inadequate and unrepresentative samplings of adults.
2) inadequacy and limitation of the materials included in the actual tests.(1)

For example, we find that most current tests of intelligence emphasize the functions which are developed in formal schooling. Therefore adults would be handicaped in them. Also not all functions slow down after middle life: there are several

---

(1) Freeman, Psychological Testing, P. 498

abilities as reasoning, vocabulary, etc... which increase for years after the age of 15 or 16. Besides, there are other factors which affect adults' performance in the test such as the different lengths of schooling, the effect of different occupations upon their mental activities and their interests and motivation which are different from those of young students. Van Wagenen states, "The hypothesis that intelligence stops maturing even earlier than the body stops growing in height is thus due partly to the nature of intelligence tests and partly to an inadequately stimulating environment."(1) In general it may be said that the mental development does not cease before the age 18-20. Psychologists agreed that mental development during the first five or six years of life is very rapid; then the amount of increase becomes smaller and smaller from year to year, until its cessation in the late teens or early twenties. Therefore they conclude for the time being that the rate and the course of mental development form a curve of negative accelaration as it is shown in Figure (2) given on the following page.

Indeed the given curve of Mental Growth is not the only form the psychologists have made about this point, but the reported curves differ in their form for many reasons: difference of age ranges in the selected samples; the various psychological processes measured by the different tests;

---

(1) M.J. Van Wagenen, "Intelligence Qotient", Encyclopedia of Modern Education, Ed. Harry N. Rivlin, The Philosophical Library of New York City, New York, 1943, P. 400.

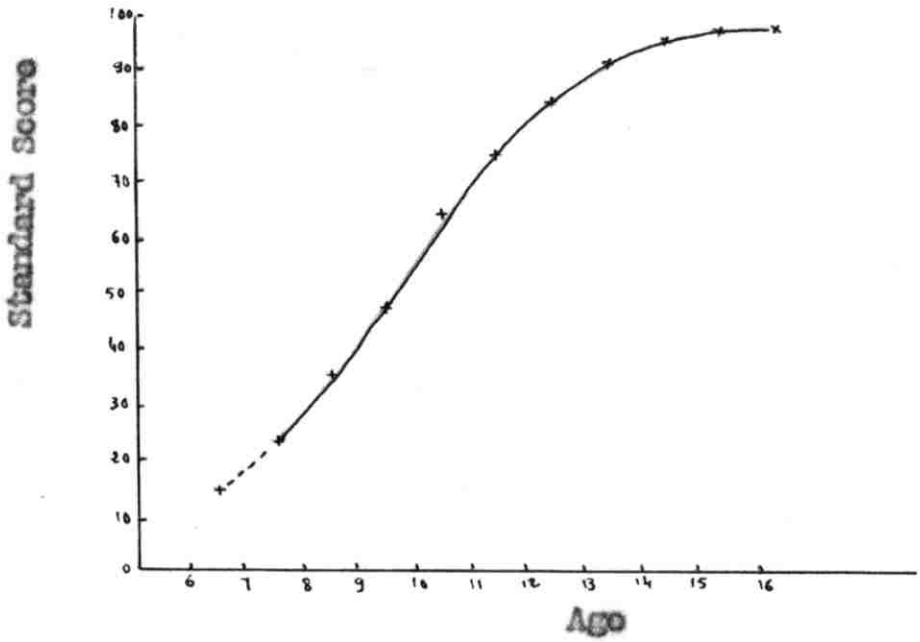


Figure (2) Curve of Mental Growth, Bellevue Full Scale, Age 7-17. (From Freeman Psychological Testing, P. 496.)

unequal validity and reliability of tests used; unequal representation of samples; the different kinds of units upon which the curves are based, etc... Moreover the growth curves representing the group trend differ from that for individuals though the former prove to be approximate to the latter.

Due to the above mentioned studies about Mental Growth, Mental Age is often substituted in the secondary school for percentile norms because the intelligence development is slow during this cycle and thus the increment from year to year is

very little. Moreover the age factor is less effective during this period in proportion to many other factors such as, social, emotional, etc..

Another aspect of Mental Growth is the observed decline in measured mental ability in late life. Data accumulated on some study shows that when mental test scores and chronological age from 15-60 years were correlated, the coefficients of correlation were - .01 to -.35 but when persons above the age of 60 were included in the samples the negative correlation increased. The question arising here is when does decline begin? and how rapid does it proceed?

In general it may be said that the growth curve of mental abilities continues as a plateau, maintaining its maximum level for some years until around the age 25-30 years then there is a very moderate rate of decline until the age of 45-50; then at last comes the rapid decline, as seen in Fig. (3) below.

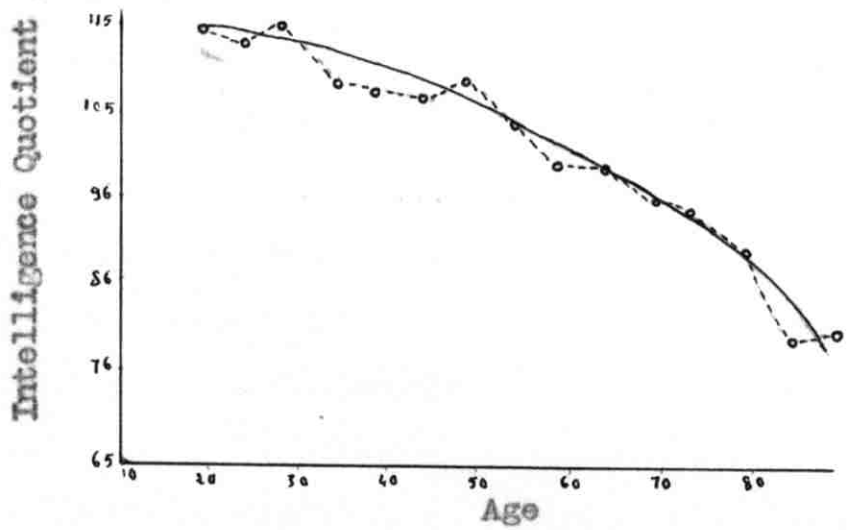


Fig. (3) Decline in IQ after Maturity. (From Freeman Psychological Testing, P. 500.

However, many points must be taken here into consideration:

1. Differences found in any one age level are greater than the differences between the different levels of age. Accordingly the highest subjects in the oldest group are better than the dullest in the younger groups.
2. Variability between individuals increases with age owing to the experiences which become more varied with age from one person to another.
3. The highly educated persons keep their superiority at all ages.
4. The curve of decline is more affected by the nature of functions measured than is the curve of growth. Some abilities increase very much with age, some a little, others decline to some extent and others a lot.
5. All intelligence tests are similar to school work. Therefore the more the individual is out of school the less he achieves in the test due to the interference of other things such as a lack of interest or exercise. Thus it has been found that the curve of scientists and scholars is much slower in its decline.
6. The actual tests are standardized on young people therefore they do not contain problems and situations that would measure the effect of experiences which are the result of age.
7. The level of education is not the same between present day old-people and the young. For the education available for a young man to-day is much higher than it was for the man of 60 or 90.

8. The samples of the old subjects are not representative since many of them who have difficulties, either with their eye sight or physical ailments, are not tested.

Until all the above-mentioned questions are settled before any curve of mental growth is plotted, we cannot decide on the real form of the curve of mental growth.

B - <u>The constancy of the IQ</u>

The question of the constancy of the IQ is very important and it has been subjected to many debates and to very long studies. For, if the IQ is constant, by testing the child once with a valid test, we can predict his later mental level. But if the individual does not keep his mental status, and his intelligence changes from time to time, we are obliged then to retest him periodically.

Extended studies have attacked this problem under the head of nature and nurture. Some on one side held the hypothesis that the IQ depends upon the heredity of the individual and is constant because the heredity of the individual does not change. Thus, data accumulated in this kind of studies have shown that most correlation coefficients between two sets of intelligence quotients of the same subjects at different times, were between *.85 and .90 .

On the other hand the second group of studies held the hypothesis that the IQ is a variable trait and it depends mostly upon the quality of the individual's training and experiences which he receives. Therefore his IQ will remain constant as long as his environment remains the same.

Consequently it was found that the IQ variation for the best ranges of Binet testing, amounts to 20 points or more, in 16 percent of the cases. Stoddard states, "we have no guarantee whatever that the obtained IQ for a child will represent his mental status at any later period of his age. We observe that a child can pass from the dull category to the bright or that a bright child can become dull".(1)

Lincoln in one of his studies found that the range of change in the IQ was from a loss of 36 IQ points to a gain of 35 points. Recent researches have revealed that a moderate change in the psychological environment will change the IQ by 15-20 percent while a fundamental change will shift the IQ up to 40 percent.(2) Again Brown add, "Subnormals with IQ's around 85 have been brought up to normal and even some morons with IQ = 60 have been made normal through specific education."(3)

English and Killian summarize this position very well by saying, "We no longer can be confident that the IQ will remain approximately constant when the environment is radically improved".(4)

The second major exception to the constancy of the IQ, other than the change in environment, is the preschool

---

(1) Stoddard, "The Meaning of Intelligence", op. cit., P. 222.

(2) J.F. Brown, Psychodynamics of Abnormal Behavior, McGraw-Hill Book, New York, 1940, P. 17.

(3) Ibid, P. 18.

(4) op. cit., Encyclopedia of Educational Research, P. 625.

tests, which are often of little value in predicting later intelligence in adulthood; especially those taken before 18 months of age are completely useless.[1] However it cannot be said that this depends entirely on the unreliability of these tests, because many of them have a fairly high degree of reliability. But other reasons may contribute to this:

1. It may be that the individual's development is more affected by environmental factors at earlier ages than later. Stoddard adds that the growth in intelligent behavior is more variable in young children, "Thus it happens occaionally that a two year old child may pass the test up through 3 and a half years getting in this way an IQ of 175. While going up in the age scale it becomes increasingly impossible for him to get such an IQ. Hence an IQ of 150 is one thing for a five year old and quite another for a 15 year old."[2]
2. The traits or functions measured by preschool tests are different from those covered by advanced tests.
3. Cattell concludes from her materials that this is "The result of changes in the tempo of development rather than the inadequacy of the tests".[3]

---

(1) op. cit., Anastasi & Foley, Differential Psy., P. 293

(2) Stoddard, "The Meaning of Intelligence", P. 219

(3) Ibid, P. 218.

Anderson tries to explain the consistancy of the IQ on the basis of overlap of abilities at successive years of a person's age. We quote him, "We deal here with a phenomenon in which the prediction of final status is based upon a larger and larger proportion of that which is included in the total; that is, scores at 10 years include more of that which is present at 16 years than do scores at 3 years .... since the growing individual does not lose what he already has, the constancy of the IQ is in large measure a matter of the part-whole or overlap relation".(1)

According to him, prediction made at age 15 is more accurate in telling an individual's future behavior than prediction at age two, because as the individual grows older, larger parts of his final status are formed.

Roff follows the same trend in explaining the constancy of the IQ. He correlated the intelligence test scores of children at different ages with their gain after one or more years and he found the correlations to be zero.(2) According to these views the increasing consistancy of mental ability is not an indication of constant rate of growth but because the present level of an individual always forms an increasing part of his level in future.

Wheeler opposes the idea of consistancy by saying, "This assertion violates a basic scientific principle that

---

(1) Anderson, "The Prediction of Terminal Intelligence from Infant & Preschool Tests" P. 388. The Thirty-Ninth Year Book of the National Society for the Study of Education Vol.XXXIX Part I, Public School Publishing Company, Bloomington, Illinois, 1940.

(2) Anastasi & Foley, Differential Psychology, P. 295.

any event is a product of its conditions and conditions can be changed".(1)

Thus we may conclude this study by saying that as long as the environment remains fairly constant the intellectually gifted children will keep their rank constant as adults, and the same thing with other individuals of varying mental levels. The score which an individual makes on a mental test is conditioned by his heredity, social status, and emotional bloking; and when the quotient remains relatively stable, it is because these factors have no marked change. Consequently we will be fair if we consider that the I.Q is the product of the interaction between many factors, some inherited and some environmental but the extent of the effect of each one is still a debatable and unsettled question.

However the environmental studies indeed have given a more hopeful view of the effect of education. Since every individual has a great many potentialities which are not ordinarily developed and since individuals seldom reach the real limits of their abilities, providing the children with everything they need at home and at schools, from birth upwards, there will be a great change in their mental status. Our concern, according to the writer's opinion must be in the relative achievement of every individual at any age; therefore it is better to measure the abilities of a student, from time to time rather than to use one test results over a

---

(1) George W. Hartmann - Gestalt Psychology (a survey of facts and principles) Ronald Press Company, New York, 1935, PP. 76,77

period of years so as to study the causes of retardation or weaknesses and to plan for a remedy. Terman summarizes this very well, "Whatever the relative contribution of nature and nurture to the subject's present intellectual status, the determination of the present status is of the utmost importance... for unless we know where the child is in his development there can be no possibility of predicting his training intelligently". (1)

Finally we may draw from our study that the range of students in the adapted test becomes no more a problem when we find that the age factor is not as important a factor in this period as is the psychological environment which the individual is surrounded by. And whenever his relative status is determined, the important thing is to follow him up in order to help him to develop to his best.

---

(1) Terman - Mental Measurement of Preschool Children - World Book 1931, P. 368.

II. The items:

It was found after the experiment and the items-analysis process that twenty items out of 90 must be eliminated from the test though there would still remain some other weak ones. This result has indicated to us the necessity for experiment on a much larger sample of items. For example in the test adapted in Sudan, 400 items were tried out on 460 children and only 88 were selected from them.

What we have come to then, is that translations of a foreign test with a priori adaptation are not sufficient and valid. To have a valid local test it must be based on extensive experiment and on a great number of different items taken from several tests and not only on one as we have done. Moreover presentation, whether in its wordings or materials, must be submitted to experiment before being put in its latest form, which thing we did not follow in our actual study and which might add to the low validity of the test. We quote Mr. Scott, "Faulty presentation not only reduced the score, but tended to invalidate the results .... I found that slight-alteration of presentation could sometimes increase or decrease the correct response by as much as 30 per cent".(1)

III. The time:

One of the most difficult questions that confronted the writer in this study was that of timing. What I have discovered after the experiment was that I underestimated the

---

(1) The British Journal of Educational Psychology, op. cit., P. 15.

time required for the test though I lengthened it by 10 minutes. Consequently a large proportion of the pupils failed to attempt all the items of the test. After 30 minutes, 8 students out of 142 finished, and 33 out of 142 reached the end of the test after 40 minutes or (23.2) percent of the students. Also 80 students out of 142, or 56.3 percent of the students, reached the seventieth item out of ninety items in 40 minutes. Therefore the time limit was still short. Accordingly I made a correlation between the number of items attempted and the score achieved for 31 students, and also 30 others and I found r to be around .89 which means that there was a high correlation between fast working and correct working with few exceptions.

Hence even the time cannot be guessed beforehand but it must be based on experiment. Moreover for this type of test and with this kind of sample time limits must not be too generous due to the high correlation between speed and power.

IV. The validity of the test:

Generally the validity of the test was not satisfactory (or in better terms was doubted) because first, no dependable valid criteria were available to compare with the test. Second, taking Teachers' Marks on the one hand and Brevet Averages on the other, as criteria, cannot be relied on very much due to their deficiency in many respects. Therefore we cannot fairly judge the low value of the actual coefficients of validity. Yet many other factors, in our opinion, have contributed to this low validity. Here are some of them:

1) The ineffectiveness of several items of the test.

2) The underestimated time designed for the test.

3) The unweighted averages of the examination marks taken as criteria.

4) The atmosphere in which the test was administered (in the Syrian University and by a new examiner), etc.. and until all these flaws can be corrected in the adapted test we cannot rely upon its results.

Briefly, we come to the conclusion that the writer's trial in this field was successful as far as becoming acquainted with techniques and procedures was concerned, but as a valid test it was doubtful.

BIBLIOGRAPHY

BOOKS

Anastasi & Foley, "Differential Psychology", the Macmillan Company, New York, 1949.

'Atiya, Mahmūd Hānā, "Al-Zakā wa Makayisuhū", Al-Nahdā Al-Masriia, Cairo, Egypt, 1949.

Ballard, Philip B., "The New Examiner", University of London Press Ltd., London, 1948

Brown, J.F., "Psychodynamics of Abnormal Behavior", McGraw Hill Book Co., New York, 1940.

Cronbach, Lee J., "Essentials of Psychological Testing", Harper & Brothers, New York, 1949.

Freeman, F.S., "Mental Tests, Their History, Principles and Applications", Houghton Mifflin Co., Boston, 1939.

Freeman, F.S., "Theory and Practice of Psychological Testing", Henry Holt & Company, New York, 1950.

Garrett, Henry E., "Statistics in Psychology and Education", Longmans Green and Co., New York, 1935.

Garrett, H. & Schneck, "Psychological Tests, Methods and Results", Harper & Brothers, New York, 1933.

Greene, Harry A., Jorgenson Albert N. & Gerberick Raymond, "Measurement and Evaluation in the Secondary School", Longmans, Green, New York, 1943.

Hartmann, George W., "Gestalt Psychology", (A Survey of Facts and Principles), Ronald Press Company, New York, 1935.

Johnson, Palmer O., "Statistical Methods in Research", Prentice-Hall Inc., New York, 1949.

Kandel, L., "Examinations and Their Substitutes in the United States", The Carnegie Foundation for the Advancement of Teaching, New York, 1936.

McNemer, Quinn., "Psychological Statistics", John Wiley and Sons Inc., New York, 1948.

Morton, R.L., "Laboratory Exercises in Educational Statistics", Silver, Burdett & Company, New York, 1928.

Peters & Woorhis Van, "Statistical Procedures and Their Mathematical Bases", McGraw Hill Book Co., New York, 1940.

Pintner, R., "Intelligence Testing Methods and Results", Henry Holt and Company, New York, 1923.

Ross, C.C., "Measurement in Today's Schools", Prentice-Hall Inc., New York, 1941.

Rugg, Harold Ordway, "A Primer of Graphics and Statistics for Teachers", Houghton Mifflin Company, Cambridge, 1925.

Rugg, Harold Ordway, "Statistical Methods Applied to Education", Houghton Company, Boston, 1917.

Sorenson, H., "Statistics for Students of Psychology and Education", McGraw Hill Company, New York, 1936.

Spearman, C.E., "The Abilities of Man", Macmillan and Company, London, 1927.

Spearman, C.E., "The Nature of Intelligence and the Principles of Cognition", Macmillan and Company, London, 1927.

Stoddard, George D., "The Meaning of Intelligence", The Macmillan Company, New York, 1947.

Terman, L., "The Measurement of Intelligence", Houghton Mifflin Company, New York, 1916.

Thorndike, E.L., Bergman, E.O., Cobb, M.V., Woodyard, Ella, "The Measurement of Intelligence", Bureau of Publications, Teachers College, Columbia University, New York.

Traxler, Arthur E., "Techniques of Guidance", Harper and Brothers, New York, 1945.

Walker, Helen M., "Elementary Statistical Methods", Henry Holt and Company, New York, 1943.

Wechsler, David, "The Measurement of Adult Intelligence", The Williams and Wilkins Company, Baltimore, 1944.

Wheeler, H. Raymond, "The Science of Psychology", Thomas Y. Crowell Company, New York, 1940.

ENCYCLOPEDIAS & YEARBOOKS

"Encyclopedia of Educational Research", Ed. Walter S. Monroe, 1941.

"Encyclopedia of Modern Education", Ed. Harry N. Rivlin, 1943.

"The Yearbook of Education", Evans Brothers Ltd., London, 1935.

"The Thirty-Ninth Yearbook of the National Society for the Study of Education", Ed. Guy Montrose Whipple., Vol. XXXIX, Part I, Public School Publishing Company, Bloomington, Illinois, 1940.

"The 1940 Mental Measurements Yearbook", Rutgers University Press, New York, 1940.

"The Third Mental Measurements Yearbook", Rutgers University Press, New York, 1949.

"The Yearbook of Education", Evans Brothers Ltd., London 1949.
"World Book," 1929.

PERIODICALS

Majalat "Al Tarbiah wa al Ta'lim", Vol. IV, V, VI, April and July, 1928. Vol. XVIII, June 1929. Vol. XIX, XX, 1929.

"Amsimat Al-Imtihanat Al-'Ammah", The Department of Education, Syrian Republic, Damascus, 1945.

"The British Journal of Educational Psychology", Vol. XX, Part I, February, 1950.

"The Journal of Social Psychology", Vol. XXIV, 1946.

"Psychological Abstracts", Vol. XXV, January, 1951.

X^{odd}	Y^{even}	R_1	R_2	D	D^2
41	42	1	1	0	0
41	40	1	2	1	1
40	36	3	14.5	11.5	132.25
39	38	4	7	3	9
37	37	5	11	6	36
36	36	7	14.5	7.5	56.25
36	33	7	21.5	14.5	210.25
36	31	7	30	23	529
35	38	1205	7	5.5	30.25
35	38	12.5	7	5.5	30.25
35	36	12.5	14.5	2	4
35	36	12.5	14.5	2	4
35	33	12.5	21.5	0	0
35	32	12.5	25.5	13	169
35	32	12.5	25.5	13	169
35	25	12.5	64.5	52	2704
34	37	20.5	11	9.5	90.25
34	37	20.5	11	9.5	90.25
34	35	20.5	21.5	1	1
34	32	20.5	25.5	5	25
34	32	20.5	25.5	5	25
34	30	20.5	36.5	15.5	240.25
34	29	20.5	42	21.5	462.25
34	27	20.5	51.5	31	961
33	39	26.5	3.5	23	529
33	38	26.5	7	19.5	380.25
33	35	26.5	17.5	9	81
33	31	26.5	30	3.5	12.25
32	38	31	7	24	576
32	34	31	19	12	144
32	30	31	36	5	25
32	30	31	36	5	25
32	28	31	47	16	256
31	39	36.5	3.5	33	1089
31	31	36.5	30	6.5	42.25
31	31	36.5	30	6.5	42.25
31	31	36.5	30	6.5	42.25
31	29	36.5	42	5.5	30.25
31	27	36.5	51.5	15	225
30	33	42.5	21.5	21	441
30	29	42.5	42	.5	0.25
30	28	42.5	47	4.5	20.25
30	23	42.5	76.5	34	1156
30	22	42.5	86	43.5	1892.25
30	22	42.5	86	43.5	1892.25
29	30	47.5	36	11.5	132.25
29	30	47.5	36	11.5	132.25
29	27	47.5	51.5	4	16
29	24	47.5	70.5	23	529
28	30	54	36	18	324
28	29	54	42	12	144
28	29	54	42	12	144
28	27	54	51.5	2.5	6.25
28	26	54	57.5	3.5	12.25
28	24	54	70.5	16.5	272.25
28	24	54	70.5	16.5	272.25
28	22	54	86	32	1024
27	30	62	36	26	676
27	28	62	47	15	225
27	26	62	57.5	4.5	20.25
27	26	62	57.5	4.5	20.25
27	26	62	57.5	4.5	20.25
27	25	62	64.5	2.5	6.25
27	22	62	86	24	576
26	28	69	47	22	484
26	26	69	57.5	11.5	132.25
26	25	69	64.5	4.5	20.25
26	23	69	76.5	7.5	56.25
26	23	69	76.5	7.5	56.25
26	22	69	86	17	289
26	17	69	117.5	48.5	2352.25
25	25	76	64.5	11.5	132.25
25	22	76	86	10	100
25	21	76	96	20	400
25	20	76	103	27	729
25	20	76	103	27	729
25	19	76	109	33	1089
25	18	76	113.5	37.5	1406.25
24	35	86	17.5	68.5	4692.25
24	28	86	47	39	1521
24	26	86	57.5	28.5	812.25
24	24	86	70.5	15.5	240.25
24	23	86	76.5	9.5	90.25
24	22	86	86	0	0
24	22	86	86	0	0
24	21	86	86	10	100
24	20	86	103	17	289
24	20	86	103	17	289
24	19	86	109	23	529
24	16	86	124	38	1444
24	13	86	136.5	50.5	2550.25
23	25	96.5	64.5	32	1024
23	24	96.5	70.5	26	676
23	22	96.5	86	10.5	110.25
23	22	96.5	86	10.5	110.25
23	21	96.5	96	.5	0.25
23	20	96.5	103	6.5	42.25
23	19	96.5	109	12.5	156.25
23	16	96.5	124	27.5	756.25
22	26	106	57.5	48.5	2352.25
22	26	106	57.5	48.5	2352.25
22	25	106	64.5	41.5	1722.25
22	23	106	76.5	29.5	870.25
22	22	106	86	20	400
22	22	106	86	20	400
22	21	106	96	10	100
22	20	106	103	3	9
22	16	106	124	18	324
22	16	106	124	18	324
22	15	106	130	24	576
21	22	114.5	86	28.5	812.25
21	21	114.5	96	18.5	342.25
21	20	114.5	103	11.5	132.25
21	17	114.5	117.5	3	9
21	16	114.5	124	9.5	90.25
21	16	114.5	124	9.5	90.25
20	18	119.5	113.5	6	36
20	18	119.5	113.5	6	36
20	17	119.5	117.5	2	4
20	15	119.5	130	10.5	110.25
19	23	126	76.5	49.5	2450.25
19	21	126	96	30	900
19	21	126	96	30	900
19	19	126	109	17	289
19	18	126	113.5	12.5	156.25
19	16	126	124	2	4
19	14	126	133.5	7.5	56.25
19	12	126	138	12	144
19	10	126	140	14	196
18	19	132.5	109	23.5	552.25
18	16	132.5	124	8.5	72.25
18	16	132.5	124	8.5	72.25
18	14	132.5	133.5	1	1
16	17	135	117.5	17.5	306.25
14	15	136.5	130	6.5	42.25
14	13	136.5	136.5	0	0
13	14	138.5	133.5	5	25
13	14	138.5	133.5	5	25
12	7	140	142	2	4
10	9	141	141	0	0
6	11	142	139	3	9

$D^2 = 60411.5$

$$P = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$$

$$P = 1 - \frac{6X60411.5}{142(142^2-1)} = 1 - \frac{362469}{142(20164-1)} = .874$$

$$r_{12} = \frac{2r_{\frac{1}{2}\frac{1}{2}}}{1+r_{\frac{1}{2}\frac{1}{2}}} = \frac{2x.874}{1+.874} = .932$$ the coefficient of reliability

Table 5! The coefficient of reliability (by spearman's Ranks Difference) before the elimination of the 20 ineffective items.

Teacher's Marks in the school exam (y variable)	Test																
	18-22	23-27	28-32	33-37	38-42	43-47	48-52	53-57	58-62	63-67	68-72	fy	y'	fy'	fy'^{2}	Σx'y'+
63-67									2(24)			2	+4	8	32	24
58-62								1(6)			1(15)	2	+3	6	18	21
53-57					1(-2)	1(0)				2(16)	1(10)	5	+2	10	20	24
48-52				2(-4)		1(0)				1(4)		4	+1	4	4	0
43-47					1(0)	3(0)	1(0)	1(0)	1(0)	1(0)	1(0)	9	0	0	0	0
38-42					1(1)	1(0)		1(-2)				3	-1	-3	3	-1
33-37												0	-2	0	0	0
28-32	1(15)					1(0)		1(-6)		1(-12)		4	-3	-12	36	-3
23-27						2(0)						2	-4	-8	32	0
F X	1	0	0	2	3	9	1	4	3	5	3	31		5	145	65
X'	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5					
F X'	-5	0	0	-4	-3	0	+1	+8	+9	+20	+15	41				
F X'^{2}	25	0	0	8	3	0	1	16	27	80	75	235				
ΣX'Y'+	15	0	0	-4	-1	0	0	-2	24	8	25	65				

Test scores (X variable)

Diagram 1. Calculation of the Product Moment Coefficient of Validity Between the Test scores and Mid-year school examination of 31 science girls in the secondary school.

$$r = \frac{\frac{\Sigma x'y'}{N} - cx'cy'}{\delta' x \delta y'}$$

But $cx' = \frac{\Sigma fx'}{N} = \frac{41}{31}$

$cy' = \frac{\Sigma fy'}{N} = \frac{5}{31}$

$\frac{\Sigma x'y'}{N} = \frac{65}{31}$

$\delta x' = \sqrt{\frac{\Sigma fx'^2}{N} - cx'^2} = \sqrt{\frac{235}{31} - \left(\frac{41}{31}\right)^2}$

$\delta y' = \sqrt{\frac{\Sigma fy'^2}{N} - cy'^2} = \sqrt{\frac{145}{31} - \left(\frac{5}{31}\right)^2}$

Therefore $r = \frac{\frac{65}{31} - \frac{41}{31} \times \frac{5}{31}}{\sqrt{\frac{235}{31} - \left(\frac{41}{31}\right)^2}\sqrt{\frac{145}{31} - \left(\frac{5}{31}\right)^2}} = .36$

Teacher's Marks in the School exam (y variable)	29-31	32-34	35-37	38-40	41-43	44-46	47-49	50-52	53-55	56-58	59-61	62-64	65-67	68-70	71-73	fy	y'	fy'	fy'^{2}	∑x'y'
68-70									$1^{(6)}$							1	+6	6	36	6
65-67				$1^{(-20)}$												1	+5	5	25	-20
62-64																0	+4	0	0	0
59-61							$1^{(-3)}$								$1^{(21)}$	2	+3	6	18	18
56-58										$1^{(4)}$	$1^{(6)}$					2	+2	4	8	10
53-55			$1^{(-5)}$		$1^{(-3)}$		$1^{(-1)}$			$1^{(2)}$						4	+1	4	4	-7
50-52	$1^{(0)}$	$1^{(0)}$	$1^{(0)}$			$2^{(0)}$							$1^{(0)}$			6	0	0	0	0
47-49		$1^{(6)}$				$2^{(4)}$		$2^{(0)}$								5	-1	-5	5	10
44-46				$1^{(8)}$		$1^{(4)}$	$1^{(2)}$	$1^{(0)}$			$1^{(-6)}$					5	-2	-10	20	8
41-43		$1^{(18)}$		$1^{(12)}$			$1^{(3)}$	$1^{(0)}$	$1^{(-3)}$							5	-3	-15	45	30
38-40					$1^{(12)}$	$2^{(16)}$		$1^{(0)}$			$1^{(-12)}$			$1^{(-24)}$		6	-4	-24	96	-8
35-37	$1^{(35)}$							$1^{(0)}$								2	-5	-10	50	35
32-34																0	-6	0	0	0
29-31					$1^{(21)}$											1	-7	-7	49	21
fx	2	3	2	3	3	7	4	6	2	2	3	0	1	1	1	40		46	356	103
x'	-7	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7					
fx'	-14	-18	-10	-12	-9	-14	-4	0	+2	4	9	0	5	6	7	-48				
fx'^{2}	98	108	50	48	27	28	4	0	2	8	27	0	25	36	49	510				
∑x'y'	35	24	-5	0	30	24	1	0	3	6	-12	0	0	-24	21	103				

Test scores (X variable)

$$r = \frac{\frac{103}{40} - \frac{(-48)}{40} \times \frac{(-46)}{40}}{\sqrt{\frac{510}{40} - \left(\frac{-48}{40}\right)^2}\sqrt{\frac{356}{40} - \left(\frac{46}{40}\right)^2}} = .12$$

Diagram 2. Calculation of the Product Moment Coefficient of Validity between test scores and. Mid-year school examination of 40 Arts girls in the secondary school.

Teachers, Marks in the School exam (y variable)	18-22	23-27	28-32	33-37	38-42	43-47	48-52	53-57	58-62	63-67	68-72	73-77	78-82	83-87	fy	y'	fy'	fy'^{2}	Σx'y'
83-87									1(6)						1	6	6	36	6
78-82															0	5	0	0	0
73-77															0	4	0	0	0
68-72									1(3)				1(15)		2	3	6	18	18
63-67					1(-6)			1(0)		1(4)	1(6)				4	2	8	16	4
58-62	1(-7)								1(1)		1(3)				3	1	3	3	-3
53-57			1(0)					1(0)			1(0)				3	0	0	0	0
48-52				1(4)			1(1)		1(-1)						3	-1	-3	3	4
43-47					1(6)	2(8)	1(2)		3(-6)	1(-4)	1(-6)			1(-12)	10	-2	-20	40	-12
38-42				1(12)	1(9)					1(-6)					3	-3	-9	27	15
33-37					1(12)		2(8)								3	-4	-12	48	20
28-32					1(15)										1	-5	-5	25	15
23-27		1(36)					1(6)				1(-18)				3	-6	-18	108	24
18-22							1(7)								1	-7	-7	49	7
F X	1	1	1	2	5	2	6	2	7	3	5	0	1	1	37		-51	373	98
X'	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6					
F X '	-7	-6	-5	-8	-15	-4	-6	0	7	6	15	0	5	6	-12				
f X '^{2}	49	36	25	32	45	8	6	0	7	12	45	0	25	36	326				
ΣX'Y'	-7	36	0	16	36	8	24	0	3	-6	-15	0	15	-12	98				

Test scores (X variable)

$$r = \frac{\frac{98}{37} - \left(\frac{-12}{37}\right)\left(\frac{-51}{37}\right)}{\sqrt{\frac{326}{37} - \left(\frac{-12}{37}\right)^2}\sqrt{\frac{373}{37} - \left(\frac{-51}{37}\right)^2}} = .26$$

Diagrams 3. Calculation of the Product Moment Coefficient of Validity between the Test scores and Mid-year school examination of 37 Science boys in the secondary school .

Teachers' Marks (y variable)	13–17	18–22	23–27	28–32	33–37	38–42	43–47	48–52	53–57	58–62	63–67	68–72	73–77	fy	y'	fy'	fy'²	fx'y'
78–82												1(25)		1	+5	5	25	25
73–77														0	+4	0	0	0
68–72												1(15)	2(36)	3	+3	9	27	51
63–67												1(10)		1	+2	2	4	10
58–62	1(−6)						1(0)	1(1)			1(4)			4	+1	4	4	−1
53–57							1(0)	1(0)	1(0)	1(0)				4	0	0	0	0
48–52			1(4)		1(2)		1(0)	1(−1)	2(−4)	1(−3)			2(−12)	9	−1	−9	9	−14
43–47			1(8)		2(8)	3(6)		1(−2)						7	−2	−14	28	20
38–42				1(9)						1(−9)				2	−3	−6	18	0
33–37														0	−4	0	0	0
28–32					1(10)									1	−5	−5	25	10
f X	1	0	2	1	4	3	3	4	3	3	1	3	4	32		−14	140	101
X'	−6	−5	−4	−3	−2	−1	0	+1	+2	+3	+4	+5	+6					
f X'	−6	0	−8	−3	−8	−3	0	4	6	9	4	15	24	34				
f X'²	36	0	32	9	16	3	0	4	12	27	16	75	144	374				
fX' Y'	−6	0	12	9	20	6	0	−2	−4	−12	4	50	24	101				

Test Scores (X Variable)

$$r = \frac{\frac{101}{32} - \frac{34}{32} \times \frac{-14}{32}}{\sqrt{\frac{374}{32} - \left(\frac{34}{32}\right)^2}\sqrt{\frac{140}{32} - \left(\frac{14}{32}\right)^2}} = .54$$

Diagram 4. Calculation of the Product Moment Coefficient of Validity between the Test Scores and mid-year school Examination of 32 Arts boys in the Secondary School.

Brevet Averages (y variable)	18-22	23-27	28-32	33-37	38-42	43-47	48-52	53-57	58-62	63-67	68-72	fy	y'	fy'	fy'^2	$\hat{\epsilon}fx'$
470-479					1(-5)							1	5	5	25	-5
460-469												0	4	0	0	0
450-459						1(0)				1(12)		2	3	6	18	12
440-449									1(6)			1	2	2	4	6
430-439						1(0)						1	1	1	1	0
420-429									1(0)			1	0	0	0	0
410-419				1(2)				2(-4)		1(-4)		4	-1	-4	4	-6
400-409						1(0)		1(-4)		2(-16)	1(-10)	5	-2	-10	20	-30
390-399				1(6)		3(0)	1(-3)					5	-3	-15	45	3
380-389					1(4)	1(0)		1(-8)	1(-12)			4	-4	-16	64	-16
370-379	1(25)				1(5)	1(0)					2(-50)	5	-5	-25	125	-20
360-369						1(0)				1(-24)		2	-6	-12	72	-24
fx	1	0	0	2	3	9	1	4	3	5	3	31		-68	378	-80
x'	-5	-4	-3	-2	-1	0	1	2	3	4	5					
fx'	-5	0	0	-4	-3	0	1	8	9	20	15	41				
fx'^2	25	0	0	8	3	0	1	16	27	80	75	235				
$\hat{\epsilon}y'x'$	25	0	0	8	4	0	-3	-16	-6	-32	-60	-80				

Test scores (x variable)

Diagram 5. Calculation of the Product moment Coefficient of Validity between the test scores and Brevet Averages of 31 Science girls in the secondary school.

$$r = \frac{\frac{-80}{31} - \left(\frac{41}{31} \times \frac{-68}{31}\right)}{\sqrt{\left(\frac{235}{31} - \left(\frac{41}{31}\right)^2\right)\left(\frac{378}{31} - \left(\frac{-68}{31}\right)^2\right)}} = .04$$

Brevet Averages	29-31	32-34	35-37	38-40	41-43	44-46	47-49	50-52	53-55	56-58	59-61	62-64	65-67	68-70	71-73	fy	y'	fy'	fy'^{2}	Σx'y'
500-509									1(7)							1	7	7	49	7
490-499				1(-24)												1	6	6	36	-24
480-489																0	5	0	0	0
470-479																0	4	0	0	0
460-469																0	3	0	0	0
450-459								1(0)	1(2)		1(6)					3	2	6	12	8
440-449																0	1	0	0	0
430-439																0	0	0	0	0
420-429						1(2)	2(2)									3	-1	-3	3	4
410-419						1(4)										1	-2	-2	4	4
400-409							1(3)	1(0)								2	-3	-6	18	3
390-399					1(12)	2(16)		2(0)								5	-4	-20	80	28
380-389			2(50)	1(20)			1(5)						1(-25)			5	-5	-25	125	50
370-379	2(84)	(36) 1		1(24)	1(18)					1(-12)	(-18) 1			1(-36)		8	-6	-48	288	96
360-369		(42) 1			1(21)		1(7)	1(0)		1(-14)					1(-49)	6	-7	-42	294	7
fx	2	2	2	3	3	4	5	5	2	2	2	0	1	1	1	35	-127	909		183
x'	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7					
fx'	-14	-12	-10	-12	-9	-8	-5	0	2	4	6	0	5	6	7	-40				
fx'^{2}	98	72	50	48	27	16	5	0	2	8	18	0	25	36	49	454				
Σx'y'	84	78	50	20	51	22	17	0	9	-26	-12	0	-25	-36	-49	183				

$$r = \frac{\frac{183}{35} - \left(\frac{-40}{35} \times \frac{-127}{35}\right)}{\sqrt{\frac{454}{35} - \left(\frac{-40}{35}\right)^2}\sqrt{\frac{909}{35} - \left(\frac{-127}{35}\right)^2}} = .08$$

Test Scores (X variable)

Diagram 6 . Calculation of the Product Moment Coefficient of Validity between the Test scores and Brevet Averages of 35 Arts girls in the secondary school .

Brevet Averages (y variable)	18–22	23–27	28–32	33–37	38–42	43–47	48–52	53–57	58–62	63–67	68–72	73–77	78–82	83–87	fy	y'	fy'	fy'²	Σx'y'
520–529													1(40)		1	8	8	64	40
510–519															0	7	0	0	0
500–509															0	6	0	0	0
490–499										1(10)					1	5	5	25	10
480–489														1(24)	1	4	4	16	24
470–479															0	3	0	0	0
460–469				1(-8)			1(-2)				1(6)				3	2	6	12	-4
450–459											1(3)				1	1	1	1	3
440–449									1(0)						1	0	0	0	0
430–439					1(3)			1(0)							2	-1	-2	2	3
420–429	1(14)				1(6)				1(-2)	1(-4)					4	-2	-8	16	14
410–419									1(-3)		1(-9)				2	-3	-6	18	-12
400–409								1(0)	1(-4)						2	-4	-8	32	-4
390–399															0	-5	0	0	0
380–389					2(+36)	1(12)	1(6)				1(-18)				5	-6	-30	180	36
370–379			1(35)			1(14)	2(14)		1(-7)						5	-7	-35	245	56
360–369		1(48)					1(8)								2	-8	-16	128	56
f x	1	1	1	1	4	2	5	2	5	2	4	0	1	1	30		-81	739	222
x'	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6					
f x'²	49	36	25	16	36	8	5	0	5	8	36	0	25	36	285				
f x'	-7	-6	-5	-4	-12	-4	-5	0	5	4	12	0	5	6	-11				
Σx'y'	14	48	35	-8	45	26	26	0	-16	+6	-18	0	40	24	222				

Tests Scores (x variable)

$$r = \frac{\frac{222}{30} - \left(\frac{-11}{30} \times \frac{-81}{30}\right)}{\sqrt{\frac{285}{30} - \left(\frac{-15}{30}\right)^2}\sqrt{\frac{739}{30} - \left(\frac{-81}{30}\right)^2}} = .50$$

Diagram 7. Calculation of the product moment coefficient of Validity between the test scores and brevet averages of 30 science boys in the secondary school .

Brevet Averages (Y variable)	13-17	18-22	23-27	28-32	33-37	38-42	43-47	48-52	53-57	58-62	63-67	68-72	73-77	fy	y'	fy'	fy'2	Σx'y'
540-549												1(56)	1(70)	2	14	28	392	126
530-539														0	13	0	0	
520-529														0	12	0	0	
510-519														0	11	0	0	
500-509														0	10	0	0	
490-499												1(36)		1	9	9	81	36
480-489														0	8	0	0	
470-479														0	7	0	0	
460-469											1(18)			1	6	6	36	18
450-459							2(-10)							2	5	10	50	-10
440-449						1(-8)							1(20)	2	4	8	32	12
430-439														0	3	0	0	
420-429								2(0)					1(10)	3	2	6	12	10
410-419														0	1	0	0	
400-409	1(0)				2(0)		1(0)		1(0)	1(0)				6	0	0	0	
390-399									1(-1)	1(-2)				2	-1	-2	2	-3
380-389						1(4)			1(-2)				1(-10)	3	-2	-6	12	-8
370-379			1(15)	1(12)	2(18)			1(0)						5	-3	-15	45	45
360-369			1(20)			1(8)				1(-8)		1(16-)		4	-4	-16	64	4
fx	1	0	2	1	4	3	3	3	3	3	1	3	4	31		28	726	230
x'	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5					
fx'	-7	0	-10	-4	-12	-6	-3	0	3	6	3	12	20	2				
fx'2	49	0	50	16	36	12	3	0	3	12	9	48	100	338				
Σx'y'			35	12	18	4	-10		-3	-10	18	76	90	230				

Test scores (x variable)

$$r = \frac{\frac{230}{31} - \frac{2}{31} \times \frac{28}{31}}{\sqrt{\frac{338}{31} - \frac{4}{(31)^2}}\sqrt{\frac{726}{31} - \left(\frac{28}{31}\right)^2}} = .46$$

Diagram. 8. Calculation of the Product Moment Coefficient of Validity Between the test scores and Brevet Averages of 31 Arts boys inthe secondary school.

Teachers' Marks in the Mid-year school exam. (y variable)	360-369	370-379	380-389	390-399	400-409	410-419	420-429	430-439	440-449	450-459	460-469	470-479	fy	y'	fy'	fy'²	ξy'x'
63-67			1(-16)				1(0)						2	4	8	32	-16
58-62					2(-12)								2	3	6	18	-12
53-57		1(-10)		1(-6)	1(-4)	1(-2)						1(10)	5	2	10	20	-12
48-52				1(-3)	1(-2)	1(-1)				1(3)			4	1	4	4	-3
43-47	1(0)	2(0)		2(0)	1(0)	1(0)			1(0)	1(0)			9	0	0	0	0
38-42			1(4)			1(1)		1(1)					3	-1	-3	3	4
33-37													0	-2	0	0	0
28-32	2(36)	1(15)	1(12)										4	-3	-12	36	63
23-27			1(16)	1(12)									2	-4	-8	32	28
F X.	3	4	4	5	5	4	1	1	1	2.	0	1	31		5	145	52
X'	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5					
F X'	-18	-20	-16	-15	-10	-4	0	1	2	6	0	5	-69				
F X'²	108	100	64	45	20	4	0	1	4	18	0	25	389				
ξY'X'	36	5	16	3	-18	-2	0	-1	0	3	0	10	52				

Brevet averages (X variable)

Diagram 9. Calculation of the Product Moment Coeffidient of Validity between Mid-year School examination and Brevet Averages of 31 Science girls in the secondary school.

$$r = \frac{\frac{52}{31} - \left(\frac{-69}{31} \times \frac{5}{31}\right)}{\sqrt{\frac{389}{31} - \left(\frac{-69}{31}\right)^2}\sqrt{\frac{145}{31}\left(\frac{5}{31}\right)^2}} = .34$$

	360-369	370-379	380-389	390-399	400-409	410-419	420-429	430-439	440-449	450-459	460-469	47o-479	48o-489	49o-499	5oo-5o9	fy	y'	fy'	fy'^{2}	ξy'x'
68-70															$_1$(42)	1	6	6	36	42
65-67														$_1$(30)		1	5	5	25	30
62-64																0	4	0	0	0
59-61	$_1$(-21)									$_1$(6)						2	3	6	18	-15
56-58	$_1$(-14)	$_1$(-12)								$_1$(4)						3	2	6	12	-22
53-55			$_1$(-5)	$_1$(-4)			$_1$(-1)									3	1	3	3	-10
50-52		$_1$(0)	$_2$(0)	$_2$(0)												5	0	0	0	0
47-49		$_1$(6)			$_1$(3)	$_1$(2)				$_1$(-2)						4	-1	-4	4	9
44-46	$_1$(14)	$_1$(12)					$_2$(4)									4	-2	-8	16	30
41-43	$_2$(42)		$_1$(15)	$_1$(12)						$_1$(-6)						5	-3	-15	45	63
38-40	$_1$(28)	$_2$(48)														3	-4	-12	48	76
35-37		$_1$(30)		$_1$(20)												2	-5	-10	50	50
32-34																0	-6	0	0	0
29-31		$_1$(42)														1	-7	-7	49	42
F X	6	8	4	5	1	1	3	0	0	4	0	0	0	1	1	34		-30	306	295
X'	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7					
F X!	-42	-48	-20	-20	-3	-2	-3	0	0	8	0	0	0	6	7	-117				
F X'2	294	288	100	80	9	4	3	0	0	16	0	0	0	36	49	879				
ξx'y'	49	126	10	28	3	2	3	0	0	0	0	0	0	30	42	293				

Brevet Averages.

Diagram 10. Calculation of the Product Moment Coefficient of Validity between Mid-year School examination and Brevet Averages of 34 Arts girls in the secondary school .

$$r = \frac{\frac{295}{34} - \left(\frac{-117}{34} \times \frac{-30}{34}\right)}{\sqrt{\frac{879}{34} - \left(\frac{-117}{34}\right)^2}\sqrt{\frac{306}{34} - \left(\frac{-30}{34}\right)^2}} = .52$$

Teachers' marks in the Mid-year School exam (Y variable).

Mid-year exam. (Y VARIABLE)	360-369	370-379	380-389	390-399	400-409	410-419	420-429	430-439	440-449	450-459	460-469	470-479	480-489	490-499	500-509	510-519	520-529	FY	y'	Fy'	Fy'2	ξ(Y-x')
83-87																		0	6	0	0	0
78-82																		0	5	0	0	0
73-77																		0	4	0	0	0
68-72					1(-12)												1(24)	2	3	6	18	12
63-67					1(-8)			1(-2)			1(4)			1(10)				4	2	8	16	4
58-62							1(-2)			1(1)								2	1	2	2	-1
53-57		1(0)						1(0)										2	0	0	0	0
48-52		1(7)									1(-2)							2	-1	-2	2	5
43-47		2(28)	2(24)				2(8)						1(-8)					7	-2	-14	28	52
38-42			1(18)				1(6)		1(0)									3	-3	-9	27	24
33-37		1(28)	1(24)			1(12)					1(-8)							4	-4	-16	64	56
28-32																		0	-5	0	0	0
23-27	2(96)					1(18)												3	-6	-18	108	114
18-22			1(42)															1	-7	-7	49	42
Fx	2	5	5	0	2	2	4	2	1	1	3	0	1	1	0	0	1	30		-50	314	308
x'	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8					
Fx'	-16	-35	-30	0	-8	-6	-8	-2	0	1	6	0	4	5	0	0	0	-81				
Fx'2	128	245	180	0	32	18	16	2	0	1	12	0	16	25	0	0	64	739				
ξx'y'	96	63	108	0	-20	30	12	-2	0	1	-6	0	-8	10	0	0	24	308				

BREVET AVERAGES (X VARIABLE)

$$r = \frac{\frac{308}{30} - \frac{-81}{30} \times \frac{-50}{30}}{\sqrt{\frac{739}{30} - \left(\frac{81}{30}\right)^2}\sqrt{\frac{314}{30} - \left(\frac{50}{30}\right)^2}}$$

$$r = .50$$

DIAGRAM 11 . CALCULATION OF THE PRO DUCT MOMENT COEFFICIENT OF VALI DITY BETWEEN MID- YEAR SCHOOL EXAMINATION AND BREVET AVERAGES OF 30 SCIENCE BOYS IN THE SECONDARY SCHOOL.

Teachers • Marksin the Mid-yearschool ex

	360–369	370–379	380–389	390–399	400–409	410–419	420–429	430–439	440–449	450–459	460–469	470–479	480–489	490–499	500–509	510–519	520–529	530–539	540–549	fx	x'	fx'	fx'²	∑x'y'
78–82																			1(45)	1	5	5	25	45
73–77																				0	4	0	0	0
68–72	1(-27)		1(-21)																1(27)	3	3	9	27	-21
63–67														1(8)						1	2	2	4	8
58–62		1(-8)			1(-5)		1(-3)			1(0)	1(1)									5	1	5	5	-15
53–57				1(0)	1(0)					1(0)										3	0	0	0	0
48–52	2(18)		1(7)		3(15)		2(6)		1(1)											9	1	9	9	47
43–47	1(18)	2(32)	1(14)		1(10)				1(2)											6	-2	-12	24	76
38–42		1(24)		1(18)																2	-3	-6	18	42
33–37																				0	-4	0	0	0
28–32		1(40)																		1	-5	-5	25	40
Fx	4	5	3	2	6	0	3	0	2	2	1	0	0	1	0	0	0	0	2	31		-11	137	222
x'	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9					
Fx'	-36	-40	-21	-12	-30	0	-9	0	-2	0	1	0	0	4	0	0	0	0	18	-127				
Fx'²	324	320	147	72	150	0	27	0	2	0	1	0	0	16	0	0	0	0	162	1221				
∑x'y'	9	88	0	18	20	0	3	0	3	0	1	0	0	8	0	0	0	0	72	222				

BREVET AVERAGES (Y VARIABLE)

TEACHERS MARKS IN THE MID-YEAR SCHOOL EXAM (X VARIABLE)

$$r = \frac{\frac{222}{31} - \frac{-127}{31} \times \frac{-11}{31}}{\sqrt{\frac{1221}{31} - \left(\frac{-127}{31}\right)^2}\sqrt{\frac{137}{31} - \left(\frac{-11}{31}\right)^2}}$$

DIAGRAM 12. CALCULATION OF THE PRODUCT MOMENT COEFFICIENT OF VALIDITY BETWEEN MID-YEAR SCHOOL EXAMINATION AND BREVET AVERAGES OF 31 ARTS BOYS IN THE SECONDARY SCHOOL.

تكييـــــف

مقاييس هنمـــن ـ نلســن للذكاء

للمدارس الثانويـــة

قامت بـــه بلقيـــس مـــروش

باشراف

دائرة التربية وعلم النفس في الجامعة الاميركية

١٩٥٠ ـ ١٩٥١

اسم الطالب ..... المدرسة ..... الصف ..... الفرع .....

التاريخ ..... العمر ..... الولادة سنة .....

ملاحظة : الاشارة * تدل على ان السؤال حذف بعد تدقيق نتائج التجربة .

ن د

ارشادات للطالب : قبل ان تباشر عليك ان تقرأ التمارين الثانية الاتية بتمعن لتفهم منها الطريقة التي يجب ان تتبعها عند اجابتك على الاسئلة .

التمرين ١ : تبطل الامطار .... في الشتاء . اشر الى الكلمة المناسبة للفراغ :

(١) ابدا (٢) قليلا (٣) بكثرة (٤) الحمراء (٥) بايمان

١	٢	٣	٤	٥

عليك ان تختار من الكلمات الخمسة التي اوردناها لك الكلمة المناسبة لمدلول الجملة . ان الكلمة الصالحة لهذه الجملة هي "بكثرة" التي تحمل رقم (٣) فعليك اذن ان تضع اشارة × في المربع الذى يحمل رقم (٣) في الجهة اليسرى من هذه الصفحة .

التمرين ٢ : يحب الطلاب ان يلعبوا : (١) كرة القدم (٢) كتبا

(٣) بلدا (٤) مطلقا (٥) ترابا .....

١	٢	٣	٤	٥

عليك هنا ايضا ان تضع اشارة × في المربع الذى رقمه رقم الكلمة المناسبة لمدلول هذه الجملة . ان الكلمة الموافقة هي "كرة القدم" التي رقمها (١) لذلك ضع اشارتك في المربع رقم (١)

التمرين ٣ : مشابهة هذا ▭ لهذا ▯ كمشابهة هذه ◯ الى :

(١) △ (٢) ▭ (٣) ○ (٤) ▽ (٥) ▯

١	٢	٣	٤	٥

الجواب هنا هو رقم (٣) لأن مشابهة مربع كبير لمربع اصغر منه كمشابهة دائرة كبيرة الى دائرة اصغر منها . لهذا ضع اشارة × في المربع الذى يحمل رقم (٣)

التمرين ٤ : ٤ ٦ ٨ ١٠ ...

... ١٦ . اشر الى العددين اللذين يجب وضعهما في الفراغين السابقين : (١) ١٧ و ١٨

(٢) ١١ و ١٣ (٣) ١٢ و ١٤ (٤) ١٤ و ١٥

(٥) ١١ و ١٢ .....

١	٢	٣	٤	٥

هنا يجب ان توءشر المربع رقم (٣) لأن الفرق بين كل عدد

وما قبله هو (٢) لذلك يجب ان يتلو الرقم (١٠) رقم ١٢ ثم ١٤ في الفراغين السابقين . وبما ان رقمهما هو (٣) لذلك ضع اشارتك في المربع رقم (٣)

التمرين ٥: ما هي الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى ؟

١	٢	٣	٤	٥

(١) كلب (٢) باب (٣) نافذة (٤) منزل (٥) طاولة...

ان الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى هي "كلب" ذات الرقم (١) لذلك ضع اشارتك على المربع ذى الرقم (١)

التمرين ٦: المجتهد وتحمق التلميذ بانتظام دروسه يحفظ. اذا رتبت هذه الكلمات لتشكل جملة مفيدة او ذات معنى ، ما هي الكلمة التي تأتي بعد كلمة التلميذ ؟

١	٢	٣	٤	٥

(١) وتحمق (٢) بانتظام (٣) يحفظ (٤) المجتهد

(٥) دروسه .....

هنا ايضا يجب ان ترتب الجملة فكريا دون ان تكتبها ثم انتخب الكلمة التي تأتي بعد كلمة "التلميذ" . ان الجملة المرتبة هي "يحفظ التلميذ المجتهد دروسه بانتظام وتحمق" . فكلمة "المجتهد" هي التي يجب ان تأتي بعد كلمة التلميذ لذلك يجب ان تضع اشارتك في المربع رقم (٤).

التمرين ٧: الفحم الى الحطب كالخشب الى : (١) الحديد

١	٢	٣	٤	٥

(٢) المدفأة (٣) الاشجار (٤) الماء (٥) الارض.....

ان نسبة الفحم الى الحطب هي كنسبة الخشب الى الاشجار ، لأن الفحم يستخرج من الحطب كما ان الخشب يستخرج من الاشجار . لذلك ضع اشارتك في المربع رقم (٣)

التمرين ٨: اذا ما احكم ترتيب الأحرف ف و م ظ . ماذا تشكل ؟ (١) وظيفة (٢) نظّف (٣) موظف (٤) نظام

١	٢	٣	٤	٥

(٥) منتظم .

هنا ايضا نجد ان كلمة (موظف) هي الكلمة الصحيحة التي تتشكل من الاحرف السابقة لذلك ضع اشارتك في المربع (٣)

==========

تنبيه :

عندما تريد تغيير الاشارة (×) بعد ان تكون قد وضعتها في مربع ما لا تحاول مسحها بالممحاة بل ارسم حول المربع المغلوط دائرة ثم اشر المربع الذى تجده صحيحا .

١٢٥ـ ما هي الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى ؟
(١) عائدة (٢) سعيد (٣) فاطمة (٤) سعاد (٥) غادة .....

١	٢	٣	٤	٥

١٢٤ـ أحسن الى حسن كأسوأ الى : (١) جيد جدا (٢) وسط (٣) سيء
(٤) سيء جدا (٥) الأحسن .....

١	٢	٣	٤	٥

١٢٣ـ ١ ، ٦ ، ١١ ، ١٦ ، ... ، ... ، ٣١ أشر الى العددين
اللذين يجب وضعهما في الفراغين السابقين: (١) ٢١ و ٢٦
(٢) ١٧ و ٢٥ (٣) ٢٦ و ٢٩ (٤) ٢٢ و ٢٧ (٥) ٢٠ و ٢٥ ...

١	٢	٣	٤	٥

١٢٢ـ كان المطر من النزارة بحيث .... الخروج . اشر الى الكلمة التي
تناسب الفراغ: (١) يدان (٢) يستسلم (٣) يتعذّر (٤) يجمع
(٥) ينكمش .....

١	٢	٣	٤	٥

١٢١ـ ١ ، ٤ ، ٧ ، ١٠ ، ... ، ... ، ١٩ أشر الى العددين اللذين
يجب وضعهما في الفراغين السابقين : (١) ١١ و ١٤ (٢) ١٥ و ١٧
(٣) ١٢ و ١٦ (٤) ١٣ و ١٦ (٥) ١٢ و ١٧ .....

١	٢	٣	٤	٥

١٢٠ـ اذا احكم ترتيب الاحرف م ت ا ح ل ماذا تشكل ؟ أشر الى
الكلمة الصحيحة : (١) محتمل (٢) انتحر (٣) مالت (٤) التحم
(٥) أحرم .....

١	٢	٣	٤	٥

١١٩ـ ٢٥ ، ٢٥ ، ٢١ ، ٢١ ، ١٧ ، ١٧ ، ... ، ... ، ما هما العددان
اللذان يجب وضعهما في الفراغين السابقين ؟ (١) ١٥ و ١٥
(٢) ١٧ و ١١ (٣) ١٣ و ١٣ (٤) ١٢ و ١٢ (٥) ١٧ و ١٣

١	٢	٣	٤	٥

١١٨ـ الماء الآسن هو : (١) الراكد (٢) العذب (٣) السريع
(٤) الصافي (٥) المزبد .....

١	٢	٣	٤	٥

١١٧ـ ما هي صلة ابنة أختي لوالدي ... ؟ أشر الى الكلمة المناسبة
للفراغ (١) ابنة اخته (٢) ابنة عمه (٣) حفيدته (٤) ابنة خميه
(٥) عمته .....

١	٢	٣	٤	٥

١١٦ـ ٥١٢ ، ٢٥٦ ، ١٢٨ ، ٦٤ ، ٣٢ ، ... ، ... ، أشر الى العددين اللذين
يجب وضعهما في الفراغين السابقين : (١) ٨ و ٤ (٢) ٢١ و ٣٠
(٣) ٣٢ و ٣٤ (٤) ١٦ و ٨ (٥) ٢٤ و ١٦ .....

١	٢	٣	٤	٥

١١٥ـ اذا ما احكم ترتيب الاحرف ل س ى ب ل س ماذا تشكل ؟
(١) سبيل (٢) سلاسل (٣) سوسن (٤) ياسمين (٥) سلسبيل

١	٢	٣	٤	٥

١١٤ـ لمدة رئيس خمس الجمهورية سنوات السورية ينتخب . اذا رتبت
هذه الكلمات لتشكل جملة ذات معنى اشر الى الكلمة التي تسبق
الكلمة الاخيرة : (١) سنوات (٢) ينتخب (٣) الجمهورية
(٤) خمس (٥) لمدة .....

١	٢	٣	٤	٥

13 ١٣ ـ ابنة عمي عندما صبي . ما هي صلة والدي لابنها : (١) ابن عمه
(٢) جده (٣) عم امه (٤) جد جده (٥) اخيه .....

١	٢	٣	٤	د

20 ١٤ ـ ٣٧، ٣٣، ٢٩، ٠٠٠، ٠٠٠، ١٧، ما هما العددان اللذان يجب
وضعهما في الفراغين السابقين ؟ (١) ٢٨ و ١٩ (٢) ٢٥ و ١٩
(٣) ٢٦ و ٢٠ (٤) ٢٥ و ٢١ (٥) ٢٥ و ٢٠ .....

١	٢	٣	٤	٥

48 ١٥ ـ يتساقط رذاذ المطر دون ... أشر الى الكلمة المناسبة للفراغ
(١) ابتداء (٢) محاكمة (٣) هدنة (٤) طول (٥) رطوبة

١	٢	٣	٤	د

١٦ ـ مشابهة هذه ⊙ لهذه ◯ كمشابهة هذا | الى :
أشر الى الشكل المناسب : (١) ◯ (٢) |
(٣) □ (٤) ⊣ (٥) ⊢ .....

١	٢	٣	٤	٥

35 ١٧ ـ ما هي صلة ابنة ابنة والدتي لابنتي ... أشر الى الكلمة
الصحيحة : (١) خالتها (٢) ابنة اخيها (٣) ابنة خالتها
(٤) امها (٥) ابنة اختها .....

١	٢	٣	٤	٥

* ١٨ ـ من خواص المياه المحافظة على ... ما هي الكلمة المناسبة
للفراغ : (١) اموالها (٢) وزنها (٣) برودتها (٤) مستواها
(٥) طولها .....

١	٢	٣	٤	د

١٩ ـ مشابهة هذه ⊕ الى ⊖ هذا ⊞ الى : (١) ⊎
(٢) ⊞ (٣) ⊔ (٤) ⊟ (٥) ⊓ .....

١	٢	٣	٤	د

22 ٢٠ ـ تخلصا المناطق ينزح الحر من الجبلية الى المصطافون . اذا
رتبت هذه الكلمات لتشكل جملة ذات معنى ، ما هي الكلمة
التي تأتي قبل (من) ؟ (١) ينزح (٢) الجبلية
(٣) المصطافون (٤) الى (٥) تخلصا .....

١	٢	٣	٤	د

* ٢١ ـ اذا أحكم ترتيب الاحرف ط ن م د ا ى ماذا تشكل ؟
(١) مطار (٢) قناطير (٣) طيور (٤) مناطيد (٥) اطنان

١	٢	٣	٤	د

" ٢٢ ـ "كثرة الطباخين تحرق الطعام" ماذا يقابل هذا المثل من
الجمل الاتية : (١) النار القوية تحرق الطعام (٢) كل
طباخ يمدح طبخه (٣) قائدان بمركب واحد يغرقانه
(٤) لا يستطيع الانسان المتمدن الحياة دون طباخين
(٥) لا يعجز القوم اذا تعاونوا .....

١	٢	٣	٤	د

٢٣ ـ مشابهة هذا (لهذا) كمشابهة هذا ⊐
الى : (١) ◡ (٢) ⊏ (٣) ⌐
(٤) ⊓ (٥) (.....

١	٢	٣	٤	د

٢٤٩ ـ ٥، ٩، ١٣، ١٧، ٢١، ٢٥، ...، ... ما هما العددان اللذان
يجب وضعهما في الفراغين السابقين : (١) ٢٩ و ٣٠
(٢) ٢٩ و ٣١ (٣) ٢٩ و ٣٣ (٤) ٢٥ و ٢٧ (٥) ٢٧ و ٢٩ ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٢٥* ـ مشابهة هذا ⊔⊔ الى ⊔ كهذا ⊔⊔⊔⊔ الى ؟
(١) ⊔ (٢) ⊔⊔ (٣) ⊔⊔⊔ (٤) ⊔ (٥) ⊔ ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٢٦ ـ الصدق الى البهتان كالتواضع الى : (١) الخوف
(٢) الجريمة (٣) الشرف (٤) التعجرف (٥) الصدق ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٢٧ ـ ٢٤٠، ١٢٠، ٦٠، ٣٠، ...، أشر الى العدد الذى يجب ان
يملأ الفراغ: (١) ٢٠ (٢) ١٥ (٣) ١٠ (٤) ٢٥ (٥) ٤٠ ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٢٨* ـ مشابهة هذا ▯ لهذا ▭ كمشابهة هذا ⬭ الى :
(١) ▭ (٢) ○ (٣) ⊙ (٤) ▯ (٥) ⬭ | ١ | ٢ | ٣ | ٤ | ٥ |

٢٩ ـ ما هي الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى ؟
(١) مصحات (٢) كتائب (٣) مدارس (٤) كليات (٥) جامعات ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٠* ـ ٩٨، ٨٧، ٧٦، ...، ...، ٤٣ أشر الى العددين اللذين
يجب وضعهما في الفراغين السابقين : (١) ٦٦ و ٤٤
(٢) ٦٥ و ٥٣ (٣) ٦٥ و ٥٤ (٤) ٥٠ و ٤٥ (٥) ٦٠ و ٥٠ ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣١ ـ الثلج الى الماء كالماء الى : (١) البخار (٢) البرد
(٣) النهر (٤) العطش (٥) الارض ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٢ ـ مشابهة هذا ▱ لهذا ▱ كهذا △ الى :
(١) ▽ (٢) ▷ (٣) ○
(٤) △ (٥) ⏢ ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٣ ـ اشر الى الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى :
(١) دخيل (٢) متواطىء (٣) شريك (٤) مساعد (٥) رفيق ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٤ ـ ٢٩، ٣٣، ٣٤، ٣٨، ٣٩، ٤٣، ٤٤، ...، ...، أشر الى
العددين اللذين يجب وضعهما في الفراغين السابقين :
(١) ٤٥ و ٤٦ (٢) ٤٥ و ٤٩ (٣) ٤٧ و ٤٨
(٤) ٤٦ و ٥٠ (٥) ٤٨ و ٤٩ ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٥* ـ أشر الى الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى :
(١) قائم (٢) موازى (٣) افقي (٤) عامي (٥) عامودى ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٦* ـ اقرب مرادف لكلمة عمام هو : (١) متأسف (٢) نعسان
(٣) شجاع (٤) سعيد (٥) نشيط ..... | ١ | ٢ | ٣ | ٤ | ٥ |

٣٧ ـ مشابهة هذا ◠ لهذا ◿ كهذه ○ الى :
(١) □ (٢) ○ (٣) ◖ (٤) ▱ (٥) ▭

١	٢	٣	٤	٥

٣٨* ـ "تمدد" هي عكس : (١) تنفس (٢) انبسط (٣) تقلص
(٤) انحجز (٥) مات .....

١	٢	٣	٤	٥

٣٩ ـ الطويل الى القصير كالنهار الى : (١) الطويل (٢) الليل
(٣) الاسبوع (٤) النهار (٥) الصباح .....

١	٢	٣	٤	٥

٤٠ ـ "الناس اعداء ما جهلوا" اشر الى اقرب معنى الى هذا المثل :
(١) الناس لبعضهم اعداء (٢) يكفيك مما ترى ما قد ترى
(٣) يستنكر الناس غالبا كل جديد (٤) ما يعتاده الانسان هو
المفيد (٥) من كان جاهلا اصبح عدوا للناس .....

١	٢	٣	٤	٥

٤١ ـ ادّخر هي عكس : (١) اراد (٢) ازدهر (٣) بدّد
(٤) وثق (٥) اصطحب .....

١	٢	٣	٤	٥

٤٢* ـ يمتاز العبقري دائما عن سواه : (١) بمواهبه (٢) باصدقائه
(٣) باختراعاته (٤) بامواله (٥) بشهرته .....

١	٢	٣	٤	٥

٤٣* ـ نسبة هذا | الى ا كهذا / الى : (١) ـــ
(٢) / (٣) | (٤) □ (٥) ـ

١	٢	٣	٤	٥

٤٤ ـ كلما تحسنت الادوات كان ذلك ... تقدم المدنية . أشر الى
الكلمة المناسبة للفراغ السابق : (١) عصر (٢) عائقا لـ ..
(٣) شؤما على (٤) دليلا على (٥) تجنب .....

١	٢	٣	٤	٥

٤٥ ـ يقصد بالمكان المنيح المكان الذي لا يمكن ان : (١) يدرك
(٢) يرى (٣) يوصف (٤) يتحدث عنه (٥) ينقل .....

١	٢	٣	٤	٥

٤٦ ـ كتاب البخلاء هو : (١) شعر (٢) تاريخ (٣) نقد (٤) فقه
(٥) سيرة .....

١	٢	٣	٤	٥

٤٧ ـ اذا كان ثمن الدفتر (٦٠) ق.س. كم دفترا يمكن ان تشترى
بمبلغ (٢٠٤٠٠) ق.س. اشر الى العدد الصحيح من الاعداد
الاتية : (١) ٣٤ (٢) ٣٤٠ (٣) ١٢٨٤ (٤) ٣٤٠٠
(٥) ١٢٨ .....

١	٢	٣	٤	٥

٤٨ ـ اشتركت الاخيرة في الامم الحرب معظم . اذا احكم ترتيب هذه
الكلمات لتشكل جملة مفيدة . ما هو اول حرف في الكلمة التي
تسبق (الامم) : (١) أ (٢) ف (٣) ح (٤) م (٥) ش

١	٢	٣	٤	٥

٤٩ ـ ٨١٠، ٢٧٠، ٩٠، ٣٠، ...، أشر الى العدد المناسب الذي يجب
وضعه في الفراغ السابق : (١) ١٠ (٢) ١٥ (٣) ٢٠ (٤) ٠ (٥) ٥

١	٢	٣	٤	٥

٥٠ ـ يستعمل الصنج في : (١) الحدائق (٢) الجراحة (٣) التصوير
(٤) النحت (٥) الموسيقى .....

١	٢	٣	٤	٥

٥١ ـ يكون الشخص التائب دائما : (١) متأسفا (٢) خالي الهموم
(٣) مريضا (٤) متباهيا (٥) محتوها .....

١	٢	٣	٤	٥

٥٢ ـ قاد الى حاكى كاخترع الى : (١) درس (٢) ابتكر
(٣) اختراع (٤) آلة (٥) عمل .....

١	٢	٣	٤	٥

٥٣ ـ غذاء الأم افضل حليب للطفل . اذا احكم ترتيب هذه الكلمات
لتشكل جملة ذات معنى ما هي الكلمة التي تسبق كلمة (الام)؟
(١) غذاء (٢) طفل (٣) حليب (٤) افضل (٥) الام - - - -

١	٢	٣	٤	٥

٥٤ ـ يحوى الدولاب دائما : (١) محورا (٢) محيط دائرة
(٣) طلاء (٤) كاوتشوك (٥) خشبا .....

١	٢	٣	٤	٥

٥٥ ـ ام ابن ابن ابي يكون بالنسبة لي : (١) ابن اخي (٢) ابن عمي
(٣) جدى (٤) اخي (٥) ابني .....

١	٢	٣	٤	٥

٥٦ ـ ٦ ، ٤ ، ٧ ، ٥ ، ٨ ، ٦ ، ٩ ، ... ، اشر الى العدد الذى يجب
وضعه في الفراغ : (١) ٧ (٢) ١٠ (٣) ٨ (٤) ٦
(٥) ١١ .....

١	٢	٣	٤	٥

٥٧ ـ اشر الى الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى :
(١) آلة (٢) اساس (٣) جهاز (٤) اختراع (٥) معدات - - - -

١	٢	٣	٤	٥

٥٨ ـ ماذا تعني كلمة (المغبة) من الكلمات الاتية : (١) المأساة
(٢) الأنفة (٣) السبب (٤) العاقبة (٥) الطمع .....

١	٢	٣	٤	٥

٥٩ ـ مشابهة هذا ⌐¬ لهذا ⌐_ كمشابهة هذا ⌐¬ الى :
(١) (٢) (٣) (٤) (٥) - - - - - -

١	٢	٣	٤	٥

٦٠ ـ اعطي لطالب ما عدد من المسائل حلّ منها بصورة صحيحة
(١٢) مسألة فكانت نسبة المسائل الصحيحة ٧٥ ٪ من عدد
المسائل كلها . فكم يكون عدد المسائل التي اخطأ فيها ؟
اشر الى الجواب الصحيح من الاعداد الاتية : (١) ٨
(٢) ٩ (٣) ٦ (٤) ٤ (٥) ٣ .....

١	٢	٣	٤	٥

٦١ ـ متى تأسست الجامعة العربية ؟
(الدقائق المجدية) ١) سنة ١٩٣٥ ٢) ١٩٣٩ ٣) ١٩٤١ ٤) ١٩٤٥ ٥) ١٩٤٩

١	٢	٣	٤	٥

٦٢ ـ اشر الى الكلمة التي لا يمكن تصنيفها مع الكلمات الاخرى (١) منصة
(٢) بناء (٣) كاتدرائية (٤) عمارة (٥) قصر .....

١	٢	٣	٤	٥

٦٣ ـ مشابهة الدائرة للشكل الاهليلجي كمشابهة المربع الى :
(الدقرلح الجديد) ١) البيضوى ٢) المكعب ٣) المنحني ٤) الدائرة ٥) المعين

١	٢	٣	٤	٥

٦٤ ـ من الّف رواية البؤساء ؟ (١) هوغو (٢) كوندورسيه
(٣) موليير (٤) شكسبير (٥) الفونس دوديه .....

١	٢	٣	٤	٥

٦٥ ـ قال نابليون ان كل جندى افرنسي يعادل (٣) جنود من النمسويين او (٥) من الروس. فكم جنديا روسيا اذن يعادل (١٢) جنديا نمسويا ؟ (١) ٢٠ (٢) ٦ (٣) ١,٥
(٤) ٦٠ (٥) ٣٦ .....

١	٢	٣	٤	٥

٦٦ ـ الشرع هو : (١) مثل (٢) قانون (٣) شراع
(٤) جدل (٥) دعوى .....

١	٢	٣	٤	٥

٦٧ ـ العلمي ولذوقه يتمتع به الذى محترما الفني لتفكيره كان .
اذا احكم ترتيب هذه الكلمات لتشكل جملة ذات معنى ما هي الكلمة التي تسبق كلمة (الفني) ؟ (١) ذوقه (٢) محترما
(٣) تفكيره (٤) الذى (٥) العلمي .....

١	٢	٣	٤	٥

٦٨ ـ* قطعة من الخشب يعادل وزنها $\frac{1}{2}$ من وزن الماء . فاذا كان الحديد يعادل (٧) امثال وزن الماء ، فكم مرة يكون الحديد اثقل من هذه القطعة ؟ انتخب الجواب الصحيح من الاعداد الاتية : (١) ٢١ مرة (٢) ٧ مرات
(٣) $\frac{1}{2}$ ٣ مرة (٤) ١٤ مرة (٥) $\frac{1}{2}$ ٧ مرة .....

١	٢	٣	٤	٥

٦٩ ـ* السرّاء للسعاده كالضرّاء الى : (١) النجاح (٢) التعاسه
(٣) اللهو (٤) الغضب (٥) الفرح .....

١	٢	٣	٤	٥

٧٠ ـ "الجلف" يعني : (١) محتقر (٢) شتّام (٣) فظ (٤) نبيل
(٥) عاقل .....

١	٢	٣	٤	٥

٧١ ـ ٣ ، ٦ ، ٨ ، ١٦ ، ١٨ ، ٣٦ ، ... ، ... ، اشر الى العددين اللذين يجب ان يشغلا الفراغين السابقين : (١) ٣٨ و ٤٠
(٢) ٣٨ و ٤٣ (٣) ٣٩ و ٤١ (٤) ٣٩ و ٧٨ (٥) ٣٨ و ٧٦

١	٢	٣	٤	٥

٧٢ ـ اذا كان ثمن ذراع القماش $\frac{1}{2}$ ٢ ليرة . فكم مترا يمكن ان نشترى بـ (١٠٠٠٠) ليرة؟ (١) ٢٥٠٠ (٢) ٤٠٠
(٣) ٤٠٠٠٠ (٤) ٢٥٠ (٥) ٤٠٠٠ .....

١	٢	٣	٤	٥

٧٣ ـ الغضب الى العنف كالمحبة الى : (١) المداعبه (٢) الكره
(٣) الاغراء (٤) الأمل (٥) السعاده .....

١	٢	٣	٤	٥

٧٤ ـ ٢٧ ، ٩ ، ٣ ، ١ ، $\frac{١}{٣}$ ، ..... ، ..... ، ما هما العددان اللذان يجب ان يملأا الفراغين السابقين ؟ (١) $\frac{٢}{٣}$ و $\frac{١}{٣}$

(٢) $\frac{٢}{٣}$ و $\frac{١}{٩}$ (٣) $\frac{١}{٩}$ و $\frac{١}{٢٧}$ (٤) $\frac{١}{٩}$ و $\frac{١}{١٨}$

(٥) $\frac{١}{٦}$ و $\frac{١}{٩}$ .....

١	٢	٣	٤	٥

٧٥ ـ تقع منطقة (التيرول) الجبلية في : (١) الألب (٢) البيرنه

(٣) الجبال الصخرية (٤) جبال لبنان (٥) هيمالايا ....

١	٢	٣	٤	٥

٧٦ ـ قيس شريط فوجد انه يتألف من ($\frac{١}{٢}$ ١٦) مترا و (٥٠) سم و (١٠٠٠) مم . فكم مترا طوله ؟ انتخب الجواب الصحيح من الاعداد الاتية : (١) ٨١ (٢) ١٩ (٣) ١٦٥

(٤) ١٨ (٥) ٣٦ .....

١	٢	٣	٤	٥

٧٧ ـ الحزن للمصيبة كالسرور الى : (١) الغم (٢) السعاده

(٣) الكره (٤) النجاح (٥) الكبرياء .....

١	٢	٣	٤	٥

٧٨ ـ "في الاعتبار غنى عن الاختبار" اى جملة من الجمل الاتية تطابق معنى هذا المثل ؟ (١) ترك ما لا يصلح اصلح

(٢) نتعلم ان نعمل بالعمل (٣) ليس الخبر كالمعاينة

(٤) الحكيم من يتعظ باخطاء غيره دون ان يقوم بها

(٥) لكل عالم هفوه .....

١	٢	٣	٤	٥

٧٩ ـ اثار الوعي العربي .... العثمانيين . أشر الى الكلمة المناسبة للفراغ : (١) كرم (٢) استعمار (٣) حفيظة

(٤) ابتهاج (٥) سيادة .....

١	٢	٣	٤	٥

٨٠ ـ يستخرج المطاط من : (١) الاشجار (٢) البحيرات

(٣) المناجم (٤) الصوف (٥) العظم .....

١	٢	٣	٤	٥

٨١ ـ مشابهة هذا لهذا كمشابهة هذا  الى :

(١)  (٢) (٣) (٤) (٥)

١	٢	٣	٤	٥

٨٢ ـ (المناوئ) هو من : (١) ناصر (٢) عادى (٣) شارك

(٤) اقتحم (٥) اعتدى .....

١	٢	٣	٤	٥

٨٣ ـ (الصوفي) هو : (١) متعبد (٢) بائع الصوف

(٣) تاجر (٤) شيخ (٥) صائف .....

١	٢	٣	٤	٥

٨٤ ـ الرجل (الوسيط) هو الذى : (١) يتدخل (٢) يحكم

(٣) يشرع (٤) يحصي (٥) يعاقب .....

١	٢	٣	٤	٥

٨٥ ــ * "أساء كاره ما عمل" ماذا يعني هذا المثل من الجمل الاتية :
(١) من سلمت سريرته سلمت علانيته (٢) كل عمل سيء يجعل صاحبه مكروها (٣) لا تسيء الى احد فتندم
(٤) من يكره على أمر لا يجيد عمله (٥) كثيرا ما يسيء الانسان ولا يدرى ما عمل .....

١	٢	٣	٤	٥

٨٦ ــ * الأمر (الجلي) هو : (١) مزعوم (٢) مكروه (٣) واضح (٤) خفي (٥) مشكوك به .....

١	٢	٣	٤	٥

٨٧ ــ * الشخص (النذل) هو : (١) صغير (٢) نزيل (٣) متأنق (٤) سعيد (٥) خسيس .....

١	٢	٣	٤	٥

٨٨ ــ نسبة التخوف الى المستقبل كنسبة الاسف الى : (١) الذكرى (٢) الامل (٣) الحزن (٤) الكره (٥) النسيان ....

١	٢	٣	٤	٥

٨٩ ــ (مستفاض) هي عكس : (١) متأكد (٢) مقتضب (٣) غامض (٤) قليل (٥) غور .....

١	٢	٣	٤	٥

٩٠ ــ ماذا تعني كلمة (الخصامي) ؟ (١) مستقيم (٢) كريم (٣) من يشرف بنفسه لا بأجداده (٤) فقير (٥) وقح

١	٢	٣	٤	٥

==========

==========

==========