INVESTIGATING THE POSSIBILITY OF USING OBJECTIVE TESTS IN PUBLIC EXAMINATIONS IN JORDAN

By

Abdallah Z. Kaylani

A Thesis

Submitted in Partial Fulfilment of the Requirements

for the Degree of Master of Arts in the

Education Department of the American

University of Beirut

Beirut, Lebanon

February, 1963

Objective Tests: Jordan

Kaylani

P R E F A C E

THE PURPOSE AND SCOPE OF THE STUDY

Public examinations in Jordan are particularly important for several reasons:

(1) At present, these examinations are held at two levels, the preparatory (intermediate secondary) and the(high) secondary. Until very recently, there have been no universities in Jordan. So, these examinations represent the end results of two terminal stages of education in the country.

(2) The relatively poor economic and social conditions have made the acquisition of a certificate one means for employment and for social mobility. The certificate acquired in a public examination has been for a long time, and to a large extent still is, used as one method of identifying those who will be active in satisfying social and economic needs.

(3) Among the large numbers of students who take the examination at the preparatory level, those who fail in the examination have no further opportunities for education in public schools.[1] Some of these students go to private schools. Others, who cannot afford to pay the fees of a private school, look for any kind of paying job, and a considerable number is left without either further schooling or employment.

The results of examinations at the secondary level are even more

---

[1] In June, 1962, about 1200 students took the examination at the preparatory level. About 75% of these succeeded.

serious. Almost all those who fail in the examination, and the greatest part of those who pass in it, are neither employed nor offered an opportunity to continue in higher education.[1] Among the graduates of this examination, only a small proportion can afford to study abroad, a certain number of those who rank at the top are offered scholarships, and a few are employed.

(4) Success and failure in these examinations have serious social and psychological implications for the students, for their parents, and for the community to which the students belong. One of the graduates of the examination at the secondary level, when asked to comment on the examination, said, "All that the examination meant to me was to get a paper they called a certificate, for this paper would enable me to continue my studies; it also pleased my parents, for failure would have meant a great distress to them; and above all I saved myself being sneered at by the people in my community." It is useful, at this point, to reproduce his criticism of the examination, "Those examinations did not test what we learned in the school. Some of the questions were irrelevant to the material in the textbook or to what teachers taught us. I used to wonder how some questions were set. How, for instance, were we, examinees, expected to answer such a vague, indefinite, and meaningless question as, Do You Think That Arabic Literature Has Fulfilled Its Aims?"

This statement made by the graduate by no means represents an

---

[1] In June, 1962, about 5000 students took the examination at the secondary level. About 75% of these succeeded.

individual attitude. The majority, if not all, of the students have the same feeling.[1]

That the results of examinations should be fair to students is of primary importance.

This study will not concern itself with the social and economic factors that lead to such mischiefs and misfits as those mentioned above.

(1) This study intends to investigate how accurate public examinations are in bringing about fair results.

There are indications that the results of public examinations include a considerable degree of inaccuracy. The apparent feature that makes one inclined to question the accuracy of these examinations is that all of them are of the essay type. A cursory reading of some of the essay-questions on the examinations demonstrates that they include a considerable degree of subjectivity and indefiniteness, and that the scoring of answers to such questions would most likely result in a considerable degree of inconsistency, and hence inaccurate scores.

There is another feature related to the accuracy of public examinations. The proclaimed policy is that public examinations are set to test the attainment of a "standard level of achievement". And here the question arises how accurately can a standard level of achievement be defined at all, and how successfully can it be then translated into an examination? While a standard level of achievement is something

---

[1] This is the writer's judgment from his experience of teaching in Jordanian schools.

theoretical which is not easy to define, in official terms it is merely a prescribed syllabus and a prescribed textbook.

(2) This study also intends to investigate how successfully public examinations test the attainment of any level of achievement, whether designated as "standard" or not.

Several studies done by research workers on testing have shown that essay examinations in general have certain serious limitations that restrict their use, as instruments of educational measurement, to certain purposes only. It was further shown, by these and other studies, that the newer, standardized objective tests are superior to essay examinations in accuracy of scoring and measurement.

Where standards cannot be easily defined, the norms provided by standardized tests can serve, at least for some time, as reasonable standards to which the standing of an individual or group on a certain function or trait that is to be assessed can be easily matched.

Standardized tests have not been tried out on any appreciable scale in Jordan, neither in public examinations nor in schools.

(3) This study intends to investigate the possibility of preparing objective tests which have the same purposes as those of public examinations, and to find out whether these tests can be more accurate and dependable measures than the essay examinations in use.

The Methods of Study:

(1) An analytical study is made of the system of public examinations as it is prescribed in official statements by authorities in the Ministry of Education.

(2) Reports of Interviews with writers of questions and with scorers of papers are presented, so as to obtain authoritative information on how questions are set and how scoring is performed, and to demonstrate the subjectivity of the questions and of scoring through the testimony of people working in the examinations.

(3) Results of Experiments done on the official scoring of standardized student answers to some of the examination questions are made to demonstrate the variability of scores assigned by different scorers to the same paper.

(4) Random samples of scores on public examinations are obtained and analyzed (statistically) to estimate the variability of scores on the examinations and the part of the variability due to human error of scoring.

(5) Objective tests on four topics of the examination at the preparatory level are prepared and administered to a sample of Jordanian students. The results are analyzed to demonstrate any superiority of measurement which can be obtained by the use of objective tests in the public examinations of Jordan.

(6) The study is supplemented at certain points by findings from other researches, especially where the study falls short in providing

the data necessary for establishing an unqualified conclusion.

Statistical treatment is provided wherever it contributes to the accuracy of formulating a result or a conclusion. The statistical principles employed are mainly those concerned with measures of variability of scores and the reliabilities of tests.

Delimitations

(1) The study is primarily concerned with the accuracy of public examinations as instruments of educational measurement.

(2) The study is confined to the system of public examinations as applied in 1962. The study will not concern itself with the changes in this system that took place in earlier years.

(3) The investigation carried out on public examinations is confined to two examinations: those prescribed at the end of the preparatory stage, and those prescribed at the end of the secondary stage.

(4) The possibility of using objective tests in public examinations is investigated for the examination at the end of the preparatory stage only. The choice of this examination for trying out objective tests was made because there were better possibilities for getting the right sample of students to whom the tests would be administered, and better opportunities for subsequent evaluation of the tests than would be the case if the examination at the end of the secondary stage had been chosen.

A B S T R A C T

Public examinations in Jordan are of the conventional essay-type, and are thus likely to exhibit the limitations common to essay examinations.

The main purpose of this thesis is to investigate the possibility of using objective tests in public examinations, and to find out whether these tests are likely to be more accurate and dependeble than the conventional essay examinations in use.

To achieve this purpose, the methods of public examinations were investigated, and objective tests were constructed and used with Jordanian students.

The investigation on public examinations included

(a) an analytical study of the present system of public examinations as it is prescribed and applied,

(b) interviews with writers of questions and scorers of papers to identify elements of subjectivity which may contribute to low validity and low reliability of the examinations,

(c) experiments on the official scoring of examinations using standardized (identical) student answers to demonstrate the variability of scores assigned by different scorers to the same student answer, and

(d) analysis of the variability of the scores on public examinations using random samples of actual examination scores for the year 1962.

The investigation on the use of objective tests consisted of preparing objective tests equivalent in content to those in the public examinations (the so called General Preparatory Examination), and trying out these tests on a sample of students who took the public examinations.

The results of the investigation on public examinations demonstrated that the subjectivity of the questions and of scoring are such as to reduce the validity and reliability of the examinations considerably.

The results of the experiment on objective tests showed that objective tests, if properly prepared and used, can be much more accurate than the essay examinations in use; the subjectivity of the questions and of scoring is eliminated, and many practical advantages are introduced.

As a result of the whole study, two proposals are made:

(1) If essay examinations are to be maintained, they can and should be improved. Suggestions for improvement are included covering those principles that should be considered when preparing essay questions and when scoring the answers to essay questions.

(2) Preliminary steps should be taken for using standardized objective tests for a substantial part of the educational measurement required. The main steps that are suggested are: (a) tryouts of objective tests at the preparatory level of public examinations; (b) using objective tests as part of the regular school methods of evaluation; and (c) providing for the initiation of testing programs that have various purposes and many useful applications.

A C K N O W L E D G M E N T S

I wish to express my deep gratitude to Professor Frederick R. Korf, the chairman of my thesis committee, whose encouragement, constant guidance, and constructive suggestions were of great help in the preparation of this thesis.

I am grateful to the members of my thesis committee, Professor Habib A. Kurani, Professor Louis P. Cajoleas, and Professor J. Edgar Lowe, for their valuable help and encouragement.

I am particularly indebted to the Ministry of Education in Amman for permitting me to conduct the investigation on the General Examinations, and to administer tests to students in its schools.

I wish to express my thanks to the writers of General Examination questions who willingly participated in the interviews on question-writing, and to the scorers of General Examination papers who participated in the interviews and the experiments on the official scoring of the Examinations. I express my appreciation to the principals of Hussein College and Hussein Secondary School in Amman for their helpful cooperation in administering objective tests to students in their schools. Appreciation is extended to the students of the first secondary class in both schools, whose responses to the tests were of great help in evaluating these tests.

Abdallah Z. Kaylani

TABLE OF CONTENTS

LIST OF TABLES

C H A P T E R 1

THE SYSTEM OF PUBLIC EXAMINATIONS IN JORDAN AT PRESENT

I. Introduction

Public Examinations and the Educational Ladder

This study is concerned with two public examinations administered by the Ministry of Education at two levels of General Education in Jordanian schools, namely at the end of the preparatory cycle and the end of the secondary cycle.

The present educational ladder of General Education consists of three cycles:

The elementary cycle - for six years,

The preparatory cycle - for three years,

The secondary cycle - for three years.

According to the latest developments in the educational system, each cycle should end with a public Government examination which serves the purpose of selecting students for the next higher cycle in the cases of the elementary and preparatory level examinations, and of providing certification qualifications for higher education in the case of the secondary level examination.[1] However, early in 1962, the

[1] Jordan, Ministry of Education, The By-Laws of General Education, No. 56 of 1960, Arts. 3 and 5. (in Arabic)

so called "Entrance Examination to the Preparatory Cycle" formerly given at the end of the elementary cycle, was eliminated. "The General Preparatory Certificate Examination" that is given at the end of the preparatory cycle was used for the first time only three years ago, in 1959. This was preliminary to the new system of secondary education that started in 1960. "The General Secondary Certificate Examination" that is held at the end of the secondary cycle has been used ever since secondary education was established in the country about thirty years ago.

Public Examinations and New Developments in the Program of General Education

In this section only those points that bear directly on the subject of study are mentioned. Requirements for admission to each cycle, the scope and orientation of the curriculum, and the organization of the educational program form the main sources for drawing certain inferences regarding the status and objectives of each of the two examinations under study.

Elementary Education:

Children are admitted to elementary classes when they complete their seventh year of age. The shortage of provisions and the crowded classes in urban areas delay admittance until seven. Authorities are planning measures to admit children of age six.

The proclaimed policy is to have the elementary curriculum

adapted to the needs of the child and of the local communities.[1] Those who set the elementary curriculum are instructed to include in it, besides a common academic core, manual and art education, agricultural education in rural schools, and domestic science in girls' schools.

Recent regulations specify that students in the first three elementary grades should be promoted without examinations.[2] Students of the sixth elementary grade who pass the school examinations are admitted to the preparatory cycle provided they are below the age of sixteen.

Preparatory Education:

For admission of students to the preparatory cycle, the By-Laws of General Education mention three criteria to be taken into consideration:[1] the availability of educational opportunities, the achievement of students in the prescribed examination, and the aptitude and academic achievement of students in the elementary classes. Since the prescribed examination was recently eliminated, the other two criteria must form the grounds for the selection of students to be admitted to the preparatory cycle. The Ministry's regulations specify an age limit for admission, a condition which is not specified in the By-Laws. And although the "aptitudes" of students are mentioned as one of the criteria for admission, we

---

[1] The By-Laws of Gen. Ed., loc. cit.

[2] Jordan, Min. of Ed., Bases of Promotion, Delinquency, and Failure; 1960, p. 1. (in Arabic)

find no prescription of how aptitudes are to be judged or used in the appraisal of students. If it is claimed that teacher-made tests measure aptitude, there is abundent evidence that almost all teacher-made tests measure almost exlusively the capacity to memorize.

The curriculum of the preparatory cycle contains a common academic core, but it is prescribed that not less than one fifth of the preparatory school program should be diversified in accordance with local needs - depending on the availability of funds and facilities.[1] Diversification includes practical instruction in Agriculture in rural preparatory schools, practical Technical and Commercial instruction in urban preparatory schools, and Home Economics in all preparatory girls' schools.

Secondary Education:

For admission of students to the secondary cycle, the By-Laws of General Education mention three criteria to be taken into consideration:[2] the availability of educational opportunities, the achievement of students in the prescribed examinations which test aptitude and academic achievement, and their scholastic grades in all classes of the preparatory cycle. In actual practice the selection of students for the secondary cycle is done exclusively on the basis of performance in a public examination known as the General Preparatory Certificate Examination, which measures academic achievement only.

---

[1] The By-Laws of Gen. Ed., loc. cit.

[2] Ibid.

The proclaimed policy urges diversification of secondary education according to the local needs and subject to the availability of funds and facilities. It further urges that the curriculum of the secondary school should be prescribed to prepare students for university work in the different areas of specialization or for life in its different walks.

Special curricula are prescribed for academic schools, agricultural schools, technical schools, commercial studies, and womens' education. Secondary agricultural schools and secondary technical schools have their own programs and their own internal examinations. Students in these schools are not required to sit for public examinations to get their graduation diplomas. Commercial study is organized in conjunction with the academic program given in academic secondary schools. The commercial program continues for three years, after which the students passing school and government examinations are granted the Commercial Certificate of the secondary level. In the academic program, division into scientific and literary sections starts in the second secondary class. Separate syllabi are prescribed for each section with emphasis on science and mathematics in the scientific section, and emphasis on languages and social studies in the literary section. The academic secondary program continues for three years, after which the students passing the government examination are awarded the General Secondary Certificate. Each student's certificate specifies the section of the Examination, Literary or Scientific, in which he passed.

II. Purposes and Objectives of Public Examinations

The statement of the purposes of an examination is important because it clarifies the status of the examination as an evaluation instrument and helps in improving both the evaluation process and the activity evaluated.

Public examinations in Jordan are organized, administered, and controlled by the Ministry of Education. Yet official statements of educational policy do not explicitly define the purposes and functions of these examinations. The prescribed System of General Examinations also does not mention "purposes" or "functions" of General Examinations. It deals mainly with administrative matters such as the organization of committees, participation requirements, bases of success and failure, etc.[1]

Observations made while the General Examinations were in progress and interviews with some members of the General Examinations-Committee have led this writer to the conclusion that "purposes" and"objectives" of public examinations are not given due consideration, and that a common-sense policy is the actual policy followed when these examinations are planned and organized. The following are samples of answers obtained from some supervisors to the question, "What do you think are the purposes of the General Examinations?":

---

[1]The System of Gen. Exams. will be discussed in section III of this chapter.

1. "To test the achievement of students."

2. "To test the attainment of a certain level of achievement which we consider as the standard - through asking about the basic concepts in the curriculum."

3. "Since we have a limited budget for education, the General Preparatory Examination serves the purpose of selecting students to the secondary cycle, where a limited number can be taken according to the availability of educational opportunities."

4. "The standard level of achievement is a primary purpose, but at present we have to consider the actual level of attainment of students - due to deficiencies in instruction and due to a common weakness in some subject areas among students."

5. "The objectives stated in the syllabus are considered the objectives of the Examination for a certain subject area."

It is to be noted that these answers represent personal opinions and not statements on which there was a general agreement or authoritative approval. Although some answers seem to contradict each other, in every answer there are some aspects which are taken into consideration in actual practice. The contradictions among statements of purposes result from external conflicting factors that influence many aspects of the educational program. Two main factors are: (1) The inadequacy of instruction which affects the setting of a standard level of achievement, (2) The inadequacy of funds and educational opportunities which reduces a claimed standard measure to a mere selecting instrument. The supposedly standard level would fluctuate with fluctuations in economic prosperity rather than with changes in the attainment of students.

In what follows, a classification of purposes of the public examinations is presented in general terms in relation to the proclaimed policy of education. To formulate these purposes in more definite terms a more careful study is needed.

1. Evaluation of educational aims:

Public examinations in Jordan could be the main methods of evaluating the attainment of educational aims. However, it is very doubtful that any study of the results of General Examinations in relation to educational aims has been made.

The general aims of education in Jordan are expressed in the following statements:

> " . . . the bringing up of a generation sound in body, belief, mind and character, which recognizes its duties toward God and the motherland, and works for the well being of the country . . ."[1]

> "The purpose of elementary education is to prepare the student to become a good citizen through developing his personality and by offering him an education aimed at forming a citizen sound in body, mind, character, and belief; it should also try to discover his tendencies and readiness for further study."[2]

> "The purpose of secondary education is to prepare the student for life and to provide him with a certain level of education that will enable the able students to specialize."[3]

The implications of these statements to the organizer of public examinations would be as follows:

(a) Public examinations should be comprehensive to the extent that

---

[1] Jordan, Min. of Ed., The Law of Education, No. 20 of 1955, Art. 3, p. 3, (in Arabic). Quotations are translated into English by this writer.

[2] Ibid, Art. 9, p. 6.

[3] Ibid, Art. 15, p. 7.

an examination can measure the various desired outcomes of education in schools.

(b) Public examinations should be capable of identifying abilities and aptitudes at different levels. They also should have discriminating power for the selection of the abler students.

(c) If examinations are to be used for identifying abilities and aptitudes and for the selection of students, they should possess reliable dispersive power.

(d) Since we aspire to bring students up to a certain level of education, examinations should measure each student's attainment relative to an explicitly defined standard level of achievement.

2. Evaluation of the objectives of the curriculum:

As the content of the curriculum and class-instruction form the medium of implementing educational aims and objectives, it becomes important for the organizer of public examinations to see how to test the realization of educational aims and objectives by testing the achievement in curriculum-content and the acquisition of desired behaiours, which may or may not have developed during the teaching-learning process.

The school-curriculum consists mainly of syllabi for subjects given in each class. Each syllabus of a subject area prescribed for a certain cycle is introduced by a set of behavioral objectives and directions for class-instruction, which are suggested by those who designed the syllabus. These objectives are essential raw material for writers of question-papers.

3. Selection of students:

Public examinations can be used, and actually are used, either to select students to a higher cycle or to choose, from among those who pass in the Examinations, the individuals who will be admitted to the various types of education that are available.

The factors that govern the selection of students are:

(a) The limited budget of education, which in turn limits educational opportunities. Selection is used to identify students who will be admitted to the preparatory cycle or to the secondary cycle - for compulsory education can be carried out in the elementary classes only.

(b) Diversification into various types of education, especially after the preparatory cycle - in secondary education. The General Preparatory Examination, if planned to meet this purpose, would be used as a basis for distributing students to the various types of secondary education. The General Secondary Examination can be planned to help students in guiding them to the kind of university specialization most suitable to them.

At present, the main criterion by which a student is selected to a certain type of secondary education if he holds the Preparatory Certificate, or to a government scholarship if he holds the Secondary Certificate is his"rank" in the respective Examination. The selection of students on the basis of their "ranks" can be misleading since a "rank" cannot be indicative enough of what

talents and abilities different students have. Moreover, the implications of a rank in the Examination affect a student's study habits and place him in an artificial situation in which he might not be able to express himself. An examination can be designed such that the different abilities of students are matched with the requirements of the different types of education that are available.

Now, in the light of what has been said, we can summarize the purposes of the two Examinations under study as follows:

The purposes of the General Preparatory Certificate Examination:

(1) To test the attainment of a standard level of achievement.

(2) To select students that will be admitted to the secondary cycle.

(3) To select, and classify, students that will be distributed to the various types of secondary education.

The purposes of the General Secondary Certificate Examination:

(1) To test the attainment of a standard level of achievement.

(2) To select students for scholarships and for Teacher Training Colleges, and to identify the best fields of specialization for students who intend to go to universities.

III. The Organization of the Examination Processes

The Ministry of Education organizes public examinations by establishing a "System of Examinations" which remains effective until it is suspended by another system. In practice, each of the General Examinations used in Jordan has its own System. The two Examination-Systems that will be considered here are: "The System of the General Secondary Certificate Examination" and "The System of the General Preparatory Certificate Examination". The two systems are similar enough to be described together, except in certain details which will be refered to in what follows. The most important items in these two Systems are the following:

1. Each of the two Examinations is held once every year - at the end of the academic year. Every participant who passes in either Examination is granted a certificate - "The General Secondary Certificate" or "The General Preparatory Certificate".

2. A Central Committee for Examinations is appointed by the Minister of Education. This Committee is responsible for executing the regulations prescribed in the Systems of Examinations. The tasks of the Central Committee are: scheduling the Examinations; asking subject-specialists to set the questions; appointing proctors and heads of the Examination-halls; assigning the places for the scoring of papers; preparing forms, publications, and circulars; and

issuing decisions and regulations related to the organization of its work. This Committee has full authority to deal with any problems that may arise during the Examination, and can take any measures to insure strict order and accurate results.

All processes of the General Secondary Examination are regulated directly by the Central Committee. In the case of the General Preparatory Examination, because of the great number of participants, the Central Committee appoints, in each of seven districts in Jordan, a "Director of the Preparatory Examination" who in turn appoints proctors, raters of papers, and other employees necessary for carrying out the administration of the Examination and producing final results. However, the same question-papers are used and the same instructions are followed in each district.

3. The General Secretary of the Examinations Committee appointed by the Minister of Education is the chief person in charge of the administrative and technical matters. He has assistants, clerks, and workers who assist him in carrying out his work.

4. Requirements for participation in the General Examinations:

I. The General Preparatory Examination is open to: (a) regular students who have studied the prescribed syllabus in a government or private school, (b) any other person, provided that not less than three years have passed since he completed elementary education in a Government or private school. Regular students are examined in the syllabus of the third preparatory class, others are

examined in the syllabi of all three preparatory classes.

II. The General Secondary Examination is open to: (a) any person who completed his study in the third secondary class in a Government or private school - he will be examined in the syllabus of the third secondary class, (b) any person who holds the Jordanian Secondary Certificate (the older system) or its equivalent - he will be examined in the syllabus of the third secondary class, (c) any person who does not have the requirements mentioned in (a) or (b) provided that not less than three years have passed since he completed the third preparatory class in a Government or private school - he will be examined in the syllabi of the second and third secondary classes, (d) any person who holds the Junior Teachers' Examination Certificate - he will be examined in the syllabi of the second and third secondary classes. In all of the above-mentioned cases it is not required that a participant should have passed the General Preparatory Certificate Examination.

The General Secondary Certificate Examination is given in four different forms to the four categories of participants:

A. Literary form given to regular students. It covers the syllabus of the third secondary class - the Literary section.

B. Literary form given to non-regular students. It covers the syllabi of the second and third secondary classes - the Literary section.

C. Scientific form given to regular students. It covers the syllabus of the third secondary class - the Scientific section.

D. Scientific form given to non-regular students. It covers the syllabi of the second and third secondary classes - the Scientific section.

Students of Commerce sit for the Examination in six subjects from the general Literary program. Their achievement in these subjects forms one part of the requirement for acquiring the"General Secondary Certificate of Commerce". Their achievement in commercial subjects in the school final examinations forms the other part of the requirements.

5. The specific subjects required, and criteria of success, in each of the two Examinations will be considered in section VI of this chapter.

IV. The Preparation of Question-Papers

The Examinations-Central Committee selects "specialized, and trust worthy" persons for writing the question-papers.[1] These persons are chosen mainly from amongst administrators and supervisors. A few of them have been chosen from among teachers in Teacher Training Colleges, but no secondary or preparatory school teacher has been asked to write Examination questions.

The setting of questions on each of the Examination subjects is entrusted to a "sub-committee" of not less than two members.[1] Actually each sub-committee for the questions of the Examinations of 1962 consisted of 3 - 4 members.

All sub-committees are called to a general meeting in which full instructions regarding the general policy, the procedure to be followed, and the responsibilities to be held are given and explained.[1] The secrecy of the questions is particularly emphasized. Thus the names of the members of the sub-committees, the dates and places of their meetings, and the dates and places of mimeographing or printing are all kept confidential, and precuations are maintained to insure security.[1] No person having close relatives (a son, daughter, brother, or sister) sitting for one of the Examinations is allowed to be a member in a subcommittee for writing questions for that

---

[1]Interview with the Assistant General Secretary of the Examinations, June, 1962.

Examination.[1]

The letter addressed to each member of the questions-subcommittees asking him to set Examination-questions specifies the following:[1]

1. A request to write Examination-questions on the subject of . . . , type of the Examination . . . , maximum grade assigned for the subject . . . , number of question-papers and the topics covered by each paper . . . , time for each paper . . . .

2. "The questions should be clear and definite so that the main points in a question can be summarized, and so that the answer to each question fits within the time allowed.

3. "The question-paper should include the following types of questions:
   a) questions related to the subject matter,
   b) questions related to the understandung of the subject matter and the student's ability to use it.
   c) as many application questions as possible.

4. "The grade assigned for each question should appear on the question-paper. The main points of the answer to every question should be prepared and submitted on a seperate paper. A fair distribution of a question's grade on the parts of the answer should be made so that the "scoring committee" will be supplied with a clearly stated grading scale.

5. "All question-writers are responsible for the secrecy of the questions. They are requested to make every precuation to maintain this secrecy so that no person who soever, and under any circumstances, is informed about the questions.

6. "The questions should be ready by . . . (two weeks from date of letter). The model answers should be submitted in a sealed envelope to the General secretary as soon as typing of the questions is completed.

---

[1]Jordan, Ministry of Education, Letter-form addressed to writers of Examination-questions, April, 1962. The letter is stamped "confidential" to conceal the name in the address. (in Arabic)

7. "You are requested to sign the attatched certificate which states that you have no son, daughter, brother, or sister in the upper classes, and who might take the Examination."

The members of each subcommittee meet and agree on a plan for organizing their work. The method of preparing questions varies from one subcommittee to another. Thus one subcommittee agrees that each member prepares one part of the questions; the complete set prepared in this way is then revised and recast in final form by the whole group.[1] Another subcommittee agrees that each member prepares a complete set of questions, the best are then selected and the final set thus prepared is revised and checked by the whole group. Still, a third group has one of its members prepare almost all the questions, the members may then revise and check these questions together.[1] There might be little criticism against these methods if revision and checking are done properly. One of the factors that determines the effectiveness of this process of revision and checking is that each subcommittee is given less than two weeks to submit the questions in final form.

The main sources of material for the questions are the syllabus and the text-book. A few members mentioned that they have used other references.[1] Some members mentioned that they were careful to include basic concepts and the essential topics of the subject in their questions; others emphasized that they tried to cover as much

---

[1] Interviews with writers of Exam-questions, June, 1962.

of the material prescribed as possible. A survey of the Examination questions of 1962 reveals that the greatest part of the questions ask about factual knowledge mentioned in the text-book.

Although instructions specified that model answers should be submitted with a fair distribution of the question's grade on the parts of the answer, many of the persons interviewed admitted that they did not do so. Some of them realized that model answers should have been prepared; others said that it was left for the scoring committee to decide upon what should be the correct answer.

All questions in both Examinations can be classified as essay-type questions. The method of structuring these questions is analyzed in the next chapter.

V. The Administration of Public Examinations

Public examinations in Jordan are administered to a relatively large number of students and other participants. To give an idea of the size of the group taking these examinations, figures from 1962 indicate that over 5,000 took the General Secondary Examination, and over 12,000 took the General Preparatory Examination. This requires the Ministry of Education to devote much effort to have the Examinations administered in a well-organized and orderly manner. Actually many of the Ministry's activities are interrupted and some of them temporarily stopped when Examinations are taking place. Some supervisors and administrators have to put their daily work aside, for some time, and take part in the Examinations. Many teachers, who by that time have their summer vacation, are employed as proctors or scorers. The work that any person offers is considered extra work and he is paid for it. All the expenses of the Examinations are collected as subscription fees from participants.

By the time Examinations start, the Examination-Committee would have made all necessary arrangements to insure smooth performance of the Examinations administration. These arrangements include the following; checking of participants' applications, preparing lists of names of examinees, assigning secret numbers to the participants, arranging for Examination-halls, appointing proctors, providing materials such as answers-booklets and stationery, preparing the

Examinations schedule, and informing participants of approval of admission to the Examinations and of regulations to be followed during the Examinations.

Special arrangements are made to keep the secret numbers key away from the knowledge of proctors, scorers, or any other person except the General Secretary and some of his assistants. This is considered an important measure for attaining sound and fair results.

The Examinations Committee circulates "Instructions for Proctoring" to every proctor before the Examinations start. These instructions form part of the Ministry's attempt to exercise strict control over public examinations. They describe in detail: measures that should be taken to maintain the security of the question-papers before these reach the hands of examinees; directions for arranging seats and examinees in Examination-halls; directions for maintaining order and smooth performance; directions for checking attendance and for reporting on misbehavior, cheating, or other incidents.

"Instructions for Examinees" are also circulated to all participants before Examinations start. These instructions describe in detail the procedure to be followed by participants while sitting for the Examinations and answering the questions. They also prescribe measures for maintaining discipline and order. It can be noted that these instructions provide examinees with clear directions for answering, which apply to all Examination-questions. It remains for examinees to think of the proper answer that satisfy the requirements set in each question. The requirements set in an essay-question, which is the type used in all papers, represents one of the problems that will be investigated in this thesis.

VI. Scoring and Final Results

Examination-papers are scored by "committees" formed from teachers, school-principals, and supervisors. Each committee scores the papers of one subject (or subject-field[1]) of one Examination level. Thus the number of committees scoring the papers of an Examination for a certain level is equal to the number of subject-fields of that Examination. In the case of the General Preparatory Examination, such committees are formed in each district (there are seven districts). Although the question-papers are the same, the final results will have independent standing for each district. In the case of the General Secondary Examination all the papers are scored by committees gathered in Amman, the capital.

The chairman of each committee organizes the work and sees that it is performed accurately and efficiently. Written instructions distributed to chairmen of committees describe the procedure that should be followed in scoring the papers. The following statement included in these instructions demonstrates the inescapable subjectivity that is expected from scoring the papers of such examinations:

> "The process of scoring is a committment and a responsibility as well. Its purpose is fair evaluation of the student's achievement and effort. Hence a successful scorer is he who makes a fair unprejudiced estimation and who considers that the future of a student depends on the paper he is scoring."[2]

---

[1] "Subject-field" refers to subjects grouped under one area of study.

[2] Jordan, Min. of Ed., Instructions to Chairmen of Scoring Committees, 1962 (in Arabic).

The procedure that is expected to be followed in scoring the papers is described in these instructions under "Bases of Scoring" in the following statements:

1. "The answer to each question should be discussed before all the members of the committee; general criteria for scoring are then set so that a question's score is carefully distributed over all parts of the answer.

2. "A subcommittee should be formed from the general committee to revise and check the scoring of others and to make sure that it is done accurately."

In actual practice, each subject-committee divides into sub-committees each consisting of 2-3 persons who are entrusted the scoring of all the papers for a particular question. A "model answer" is formulated for each question, and a "scale" for scoring each part of the answer is established. All the members of the committee participate in setting the "model answer" and the "scoring scale".[1]

The customary trend is that a person called a "corrector" reads the answers in every paper, estimates a score according to the established scale and records this score, numerically, on the

Question No.	Corrector's Score (Numerical)	Revisor's Score (in words)	Revisor's Signature	Remarks

cover of the answers-booklet in a special blank rectangle (see diagram). Another person called a "revisor" reads independently the same answer, estimates another score, and if it is the same as that

[1] Interviews with scorers of papers, June, 1962.

estimated by the "corrector", the score is approved by writing it in words in another column on the cover of the answer-booklet and by having the "revisor" sign his initials next to his score in words.

The following observations were made by the writer when he visited some of the halls where committees were engaged in scoring Examination-papers:

- All committees and subcommittees had set model answers and a scoring scale for each answer.

- In many cases, the answers to one question were scored by one person and revised by another.

- In some cases, one person revises the work of two sub-committees.

- In a few cases a person is entrusted the scoring of answers to a question, alone , without being checked.

- At the beginning of the work, revision and checking are carried out for every paper. Later, the revisor checks a proportion which varies from one subcommittee to another and from time to time. In some committees, checking is not done on more than 30% of the papers.

- Disputes about a score between a "corrector" and a "revisor" are supposed to be settled through argument and discussion. This is the actual practice when the difference between the two estimations is apparently serious. In most cases it was noticed that a revisor changes a score authoritatively without consulting his colleague.

- Revision is done at a much faster rate than original scoring. As time goes on the proportion of papers revised is reduced, and about

70% of the papers are "revised" by just writing the same grade in words and appending initials. It is not unusual to see revisors chatting with each other or drinking refreshments while correctors are immersed in their work.

Bases for Passing in the General Secondary Certificate Examination:

I. The subjects required in the General Secondary Certificate Examination, the time assigned for each subject paper, and the maximum and minimum grades prescribed for each subject are shown in tables 1 and 2.

All participants must take the Examination in all the subjects of either the Literary program or the Scientific program.

II. "Grand total of grades for each student is obtained by adding the student's grades in 8 subjects for boys and in 9 subjects for girls, provided that this total includes: (a) for literary students, the grades of Religious Education (for Moslem students), Arabic Language, and English Language; and (b) for scientific students, the grades of Religious Education (for Moslem students), Arabic Language, and Mathematics. In all cases the student's highest grades in other subjects are included in his grand total."[1]

III. "A participant passes in the Examination of the Literary program provided that (a) he passes in at least six subjects among which must be Arabic Language -for all students, and Religious

---

[1] Jordan, Min. of Ed., The System of the General Secondary Certificate Examination for 1961; (in Arabic).

Table 1

The General Secondary Certificate Examination

The Literary Program

Subject	No. of Exam. Papers and Time Assigned for Each Paper	Maximum Grade	Minimum (Passing) Grade
1. Religious Education (Compulsory for Moslem students only)	1 paper 3 hrs.	100	40
2. Arabic Language	2 papers 2X3 hrs.	300	150
3. English Language	2 papers 2X3 hrs.	300	120
4. Arabic History	1 paper 3 hrs.	100	40
5. General History	1 paper 2 hrs.	100	40
6. Arabic Geography	1 paper 2 hrs.	100	40
7. General Geography	1 paper 2 hrs.	100	40
8. General Mathematics and General Science	1 paper 3 hrs.	100	40
9. Social Science and Philosophy	1 paper 3 hrs.	100	40
10. Women's Education (for girls)	1 paper 3 hrs.	100	40

Table 2

The General Secondary Certificate Examination

The Scientific Program

Subject	No. of Exam. Papers and Time Assigned for Each Paper		Maximum Grade	Minimum (Passing) Grade
1. Religious Education (Compulsory for Moslem students only)	1 paper	3 hrs.	100	40
2. Arabic Language	2 papers	$1\frac{1}{2}$ + 2 hrs.	200	100
3. English Language	2 papers	$1\frac{1}{2}$ + 2 hrs.	200	80
4. Mathematics	2 papers	2X3 hrs.	300	120
5. Physics	1 paper	3 hrs.	100	40
6. Chemistry	1 paper	3 hrs.	100	40
7. Biology	1 paper	3 hrs	100	40
8. Specialization in one science	1 paper	3 hrs.	100	50
9. The Arab Land	1 paper	2 hrs.	100	40
10. Women's Education (for girls)	1 paper	3 hrs.	100	40

Education -for Moslem students only; (b) the sum of two grades in two subjects other than the six required in (a) is not less than 25% of the sum of the maximum grades for these two subjects.[1]

IV. "A participant passes in the Examination of the Scientific program provided that (a) he passes in at least six subjects among which must be Arabic Language -for all students, Religious Education for Moslem students only, and any three scientific subjects;[2] (b) the sum of two grades in two subjects other than those required in (a) is not less than 25% of the sum of the maximum grades for these two subjects."[1]

V. "The Examinations Committee makes a survey study of the final results and submits its recommendations to the Minister of Education before names are announced. After the Minister approves the results he announces the names of those who passed in the Examination."[1]

The recommendations made by the Examinations Committee deal with such questions as the percentage of success in the Examination, proposed changes in the passing grade of some subjects where too many failures occur, proposed decisions regarding individual cases such as the case of an individual having a failing grade close to the passing grade when his other grades are outstandingly good, etc.

Bases for Passing in the General Preparatory Certificate Examination:

I. The subjects required in the General Preparatory Certificate

---

[1] Ibid.

[2] Scientific subjects include Mathematics, Physics, Chemistry, Biology, and the subject of specialization.

Examination are the following:[1]

1. Religious Education -for Moslem students.
2. Arabic Language
3. English Language
4. Mathematics
5. Social Studies
6. Sciences
7. One of the following:
   (a) Technical Education
   (b) Commercial Education
   (c) Agricultural Education
   (d) Domestic Science (for girls)

The Examination in each subject consists of one 2-hours paper for all the subjects. All of the subjects are assigned the same maximum grade of 100 and the same passing grade of 50.[1]

II. A participant passes in the General Preparatory Examination when he passes in at least 5 subjects, among which must be Arabic Language.[1]

III. Each district Examination-committee makes a survey study of the final results in the district. It submits the results with its recommendations concerning them to the Minister of Education. The Minister approves the results after consulting with the Central Examinations Committee, and then announces the names of those who passed in the Examination.[1]

---

[1] Interview with the Assistant of the Director of the Prep. Exam. in Amman District; June, 1962.

C H A P T E R 2

ANALYSIS OF THE METHODS USED IN PUBLIC EXAMINATIONS: A RESEARCH STUDY MADE ON THE EXAMINATIONS

I. The Problem and the Methods of Research

It has already been mentioned that public examinations in Jordan are of the traditional essay-type examinations. According to modern educational theory, essay examinations have certain serious limitations. Generally they suffer, in comparison with objective tests, from low validity and low reliability. Moreover, they are not easy to administer and score and are time consuming.

A tentative hypothesis can thus be formulated, namely, that public examinations in Jordan are likely to exhibit the limitations common to essay examinations, primarily low validity and low reliability.

An investigation on the two General Examinations, the Preparatory and the Secondary, was carried out to check on the truth of this hypothesis. It is hoped that the investigation has produced convincing evidence which testifies to the ineffectiveness of conventional methods and that it may help to change the belief that the mere application of strict controls or the prescription of regulations and

instructions can result in improvement or progress, when actualities indicate that radical changes are needed; that new methods more effective than the conventional ones should be used.

The investigation was confined mainly to the Examination-questions -how they are prepared, and scored. The preparation of questions and the scoring of papers were considered to be the main factors that determine the extent to which Examinations are valid and reliable.

The methods of investigation used were:

i) Interviews with writers of Examination-questions,

ii) Interviews with scorers of Examination papers,

iii) Experiments on the official scoring of standardized student-answers,

iv) Analysis of random samples of Examination grades.

The description of these investigations and their results is presented in sections II to V of this chapter. The last section, VI, is devoted to the interpretation of results.

II. Interviews with Writers of Examination-Questions

The purposes of interviewing writers of Examination-questions were: (a) to determine the sources of subjectivity in the questions set for the Examinations, (b) to discover general tendencies in planning and constructing Examination-questions, and (c) to find out what elements in the setting of questions would have affected the validity and reliability of the Examinations.

From about fifty persons taking part in setting questions for the two General Examinations, the Preparatory and the Secondary, of 1962, only nine persons were interviewed. These nine persons were not randomly chosen, for the attempt was to have as many as possible participate in this investigation. Since most of the persons who wrote Examination questions occupied high positions and were noted to be very busy, only those persons who could offer some of their time and willingly consented to participate were interviewed.

A permission to proceed in this type of investigation was secured from the General Secretary of the Examinations. He only emphasized that all information collected should not be publicized by having it propagated to persons who might use it for personal prejudice.

The interviews were conducted in a rather informal manner, but the discussion was restricted to the issues suggested by the

questions that were presented in the interviews. Interviews were conducted in Arabic. Concise statements of answers or comments delivered during an interview were recorded directly after they were pronounced. These statements are presented in the following pages in summary form, by grouping together all the answers that belong to a particular question and then summarizing each set of answers. It must be noted here that results and conclusions are based not only on responses produced by the majority, but also on individual responses, since each interviewee can be assumed to represent a subcommittee that set questions on a subject-field which counts one eighth of the entire Examination, and therefore has a significant effect on its outcome.

The nine persons interviewed represent six subcommittees for writing questions on six subjects, namely Arabic Language, English Language, History, Social Science, Mathematics, and Biology. Four of them took part in writing questions for both the Secondary and the Preparatory Examination levels. Five of them were appointed chairmen of committees for scoring the Secondary Examination papers. One was Assistant Under Secretary of State in the Ministry of Education, seven were central inspectors, and one was a teacher in Teacher Training College who was later appointed Cultural Attchè in Turkey. One or two do not hold university degrees, but they have had training courses of university level in addition to long experience in teaching. The rest are subject specialists and hold university degrees, B.A. (4), M.A. (1), and Ph.D. (2). Some of the latter persons have

had post graduate training courses. Eight of them have had experience in inspection, but all of them have had experience in teaching and had taken part in setting the school syllabus of the subject of their specialization.

The following section outlines the interview questions and answers:

The Main Questions in the Interviews with Writers of Examination-Questions:

(1) How did you plan your questions? Was there any written plan? What did your plan consist of?

(2) What were the main sources of your questions? The syllabus? The text-book? Other books?

(3) What criteria were considered in selecting the material for your questions? What experiences were useful to you in identifying the appropriate material for the questions?

(4) What objectives were thought of as pertaining to the subject of your questions?

(5) How do you like to start your question, with "Discuss", "Explain", "Mention", "Compare", etc.? Do you have a favorite way of structuring your questions? Which type of question do you prefer to write, the general question or the one which specifies the answer?

(6) Were model answers to your question prepared and presented to the Examination Committee?

Statements Made by Question-Writers in Answer to Interview-Questions

Question (1): How did you plan your questions? Was there any written plan? What did your plan consist of?

Answer-Statement	Respondents No.	%
(a) No written plan	8	100%[1]
(b) The plan consisted of organizing the work, assigning tasks to each member, and arranging for next meetings.	6	75%
(c) Each member prepared independently "part" of the questions. The parts were then brought together and produced in final form by the participation of all the members.	5	62%
(d) Each member prepared independently a "whole set" of questions. The best questions were then selected, revised, and produced in final form by the participation of all the members.	3	38%
(e) I was the main person who planned and did the work. Some colleagues in my group did not participate systematically.	2	25%
(f) My group faced two problems: (1) no text book is prescribed for the subject and (2) teachers differed in teaching the prescribed syllabus. Our plan consisted of giving due consideration to such problems.	1	12%
(g) It is important to set a policy before writing questions for an examination.	1	

[1] In this part, 100% of respondents is taken to be 8 and not 9 because one interviewee did not take part in setting questions for the examinations of 1962, but he wrote examination-questions in previous years. Statement (g) above represents his point of view .

Question (2): What were the main sources of your questions? The syllabus? The text-book? Other books?

Answer-Statement	Respondents No.	%
(a) The syllabus	9	100%
(b) Text-books	5	55%
(c) Other sources	2	22%

Other sources mentioned were: (1) questions from previous examinations of the Egyptian General Secondary Certificate and London General Certificate and (2) some books, Arabic, English, or American, that have the subject-matter of the same level as that specified for students taking the examinations.

Question (3): What criteria were considered in selecting the material for your questions? What experiences were useful to you in identifying the appropriate material for your questions?

Answer-statement	Respondents No.	%
(a) The syllabus (the main criterion)	9	100%
(b) The essential points in the syllabus that students should know.	1	11%
(c) The essential points in the text-book.	1	11%
(d) Basic concepts and basic ideas in the syllabus.	1	11%
(e) Every item in the syllabus is important.	2	22%
(f) "The actual level of attainment of students" was cited as a second criterion by two persons who were in the same subcommittee for writing questions for the English Language Examination.[1]	2	22%

[1] It is generally acknowledged in official educational circles that students are in general weak in English. The percentage of success in English in the official examinations is usually one of the lowest percentages.

In answer to the second part of the above question, a few persons mentioned certain experiences that helped them in identifying the material appropriate for the questions. These experiences are actually common to almost all others who were on the committees for writing Examination-questions. They consist of specialization in a subject, experience in teaching and inspection, participation in setting the syllabus, and in many cases the authorship of textbooks.

One interviewee mentioned that "basic concepts and basic ideas" should be identified by the following criteria: (a) they should be related to daily living, (b) they should have significance for future study, (c) they should be related to their importance in the curriculum -that is, the time actually devoted to them, and (d) they should stress characteristics that show basic abilities, reasoning, implementation of study, etc.

Actually, this statement represents a personal opinion of what ought to be and not of what actually is taking place. The statement comes from a person whose position entitles him to be one of those who set the official policy.[1]

Question (4): What objectives were thought of as pertaining to the subject of your questions?

The answers to this question were presented in four different ways:

---

[1]This person has been recently appointed Assistant Under Secretary of State for Higher Education.

	Respondents No.	%
(a) by stating what the questions are supposed to measure,	3	33%
(b) by stating some of the objectives that occured in questions already written,	3	33%
(c) by stating that the objectives of the course that are specified in the syllabus were used,	2	22%
(d) by admitting that there was no thinking of objectives when questions were written.	1	11%

In all the answers there was no indication that well-defined objectives have been set prior to the setting of the questions or that the questions have been planned according to objectives and criteria carefully chosen.

Except for the two persons who referred to the objectives in the syllabus and thus they did not mention specifically what objectives were used, all others gave examples of objectives pertaining to their respective topics. Some of these objectives were very general , such as, "testing abilities acquired in the school", "testing for thinking", or "giving due consideration to the level of students"(!) Other objectives mentioned were rather specific, such as, "testing the understanding of the literary work and the ability to evaluate it", "testing the ability to analyze", "testing the ability to express in correct language", etc.

Some of the objectives mentioned exhibit the general tendency of some question-writers to make the questions, perhaps unconsciously, highly subjective. Examples of such objectives are:

"to test the ability to evaluate literary work",
"to test the ability to organize ideas",
"to test the ability to present sound ideas of national significance",
"to test the ability to present a good argument",
etc.

Question (5): How do you like to start your question, with "Discuss", "Explain","Mention", "Compare", etc.? Do you have a favorite way of structuring your questions? Which type of question do you prefer to write, the general question or the one which specifies the answer?

Answer-Statement	Respondents No.	Respondents %
(a) Favors no special way of structuring a question.	5	56%
(b) In favor of writing general questions in which the student is given some freedom to deal with the question in his own way.	3	33%
(c) In favor of writing definite questions that have specific answers.	3	33%
(d) Likes to start a question with phrases such as: Talk about, Explain, Discuss, Present the argument for, or State view or opinion.	5	56%

The way the above question was answered in the interviews showed that the significance of the method of structuring questions was not very clear in the minds of the persons interviewed. The majority stated that they prefer no particular way of structuring a question. Even when preferences were indicated, such preferences are likely to have been mere expressions of personal taste or opinion

rather than the expression of the understanding of the principles underlying an evaluation instrument they had been constructing.

The above results can be further illustrated by considering the questions written for the General Examinations of 1962. A general review of these questions shows that the way of structuring questions varies greatly from one paper to another but varies little among questions in the same paper. Thus in one paper we find that most of the questions start with a statement followed by "explain", "analyze", "discuss", etc. In another paper the prevailing patterns are "Talk about", "Write all you know about", "Talk in brief", etc. Some of these questions can be considered as subjective and having no definite answers. Examples of such questions are given below. Other questions tend to be less subjective and to have more definite answers. Examples of these questions are mathematical problems and some questions on grammar.

Examples of questions given in the Examinations that tend to be very subjective:[1]

- Explain the influence of the Industrial Revolution on: society, the family, the workers, work-owners, and imperialism.
- Talk about the inhabitants of Algiers; discuss the bad distribution of the land; mention the products and state their significance to the economic development which will start.
- Talk in brief about the basic ideas and meanings in the story of Hay Ben Yakzan.
- Talk about freedom of opinion in Islam.

---

[1]Literal translation by this writer of the questions from the original in Arabic.

- Discuss the circulatory system in rabbits.
- Discuss in detail the process of photosynthesis in plants.

The kind of phrase used at the beginning of each question can be used as one criterion for determining the degree of subjectivity in a question. The kinds of phrases used in the questions of public examinations are sampled below by taking the group of sets of questions on the General Secondary Examination -the Literary Section and the group of sets of questions on the General Preparatory Examination.

On the General Secondary Examination -Literary Section, there are about 85 questions covering the first 9 subjects in Table 1 page 26. The average time allowed for answering a question is about 22 minutes. The phrases used in these questions can be roughly classified as follows:

Type of Phrase	Frequency	Frequency Percentage
(a) Write about (one of the topics);[1] What do you think; What is your opinion	8	9%
(b) Talk about -in brief, in detail; Write about	13	15%
(c) Explain; Explain in brief	15	18%
(d) Discuss; Discuss in detail, in brief	7	8%
(e) Statement or opinion followed by "explain", "analyze", "why" ...	17	20%
(f) What (the definitive)	14	17%
(g) Application of rules (grammatical and mathematical)	11	13%

[1] There are three questions on composition in the whole Exam.

On the General Preparatory Examination, there are about 50 questions covering the first six subjects on page 29. The average time allowed for answering a question is about 15 minutes. The phrases used in these questions can be roughly classified as follows:

Type of Phrase	Frequency	Frequency Percentage
(a) Talk about; Write about; What did you learn from ...	8	16%
(b) Explain; Discuss; What is the importance of..	10	20%
(C) Write about one of the topics (composition); Paraphrase; Translate	5	10%
(d) What; Mention; Define; Ennumerate	11	22%
(e) Application of rules (mathematical and grammatical)	16	32%

In the above classifications, the first five categories identified in the Secondary Examination-questions and the first three categories identified in the Preparatory Examinatio-questions include types of phrases that tend to make the questions open, indefinite, and subjective. According to the rough estimations given above, such phrases occur in about 70% of the questions on the General Secondary Examination (the Literary Section) and in about 45% of the questions on the General Preparatory Examination.

It can be further noticed that about 20% of the questions on either Examination are, by virtue of using certain phrases such as "What", "Mention", "Ennumerate", etc., of the straightforward recall

type. Further investigation showed that the percentage of the recall-type question is much higher than 20%. It was testified by many scorers, when these were interviewed, that all points in the model answer were taken directly from the textbooks. Astonishingly enough, it was found that the questions which appear to call the student's special abilities, to reason, analyze, evaluate, etc. have full answers reproduced in textbooks. It can then be said that, excluding composition questions and those questions that require the application of rules, almost all questions are of the simple or selective recall type.

Question (6): Were model answers to your questions prepared and presented to the Examination Committee?

Answer-Statement	Respondents No.	Respondents %
(a) Model answers were prepared and submitted. (Mathematics, Biology, Social Science)	3	38%
(b) Model answers were not prepared. (Arabic Language)	2	25%
(c) Model answers to some questions were prepared and submitted. (English Language)	2	25%
(d) Model answers, in outline form, were prepared and submitted. (Geography)	1	12%

III. Interviews with Scorers of Examination-Papers

The purpose of interviewing scorers of Examination-papers was to investigate some of the factors that affect the decision needed when estimating a score for a certain answer. The assumption made here is that factors resulting from personal attitudes of scorers and from the nature of the structure of the essay-type question would most probably result in varying degrees of inaccuracy or inconsistency in scoring.

This inquiry is also intended to supplement the results of experiments made on the official scoring of standardized student answers, to be described in the next section.

The interviews were made at the time when the papers of the General Secondary Examination of 1962 were scored. A convenient date was chosen so that each committee had already been in session for a few days and had finished the scoring of a considerable number of papers. This was also the date on which the experiments on official scoring were performed.

The conditions set for controlling the process of scoring placed many restrictions on the investigator, and made it difficult to undertake a more extensive study than this one. Interruption of actual scoring was allowed only by special permission and then only for the shortest time possible. Each interview did not last more

than six minutes, and only fifteen scorers were interviewed. Eight of them were on the Mathematics committee of 25, and seven were on the History committee of 45. The two groups of 8 and 7 were selected from their respective committees by approximate random sampling. Scorers were distributed in rooms in the same building. The number of scorers selected for the interview from each room was approximately proportional to the number present in the room. No particular order of selection was followed as there was not one order of seating in all the rooms.

The interviews took the form of free conversation, in Arabic. Notes were taken while each person was being interviewed. English translations of summary statements of the answers to each of the five questions asked in the interview are given below. The procedure followed was similar to that outlined in the preceding section.

It may be noted that the results of the interviews are in some cases insufficient for the drawing of general conclusions. But these interviews were nevertheless useful in two ways: first, they provide real examples on some of the factors that affect the accuracy of official scoring of examination-papers; second, they suggest tentative hypotheses on issues or problems that can be more extensively investigated in the future.

<u>The Main Questions in the Interviews with Scorers of Examination-Papers</u>:

(1) What were some of the difficulties and causes of inaccuracy that you met when you were estimating scores for answers to your

questions?

Suggested examples: poor organization, poor expression and language, poor handwriting, unsatisfactory explanation, etc.

Please mention any other examples pertaining to your question.

(2) To what extent did you subject your estimation of a score to the pre-established scale? Were you influenced by some elements in the answers of students other than those mentioned in the model answer.

Suggested examples: style; clarity of thought and expression; good, or bad organization; etc.

Please mention your own examples.

(3) Did you notice from the answers you read that the question was ambiguous, misunderstood, or answered differently from what was intended?

(4) Do the scores on the answers you scored tend to be high, average, low, or varying?

(5) On what proportion of the papers does the "revisor" check the scores by going back to the answers and reading them again?

Statements Made by Scorers

In Answer to Interview-Questions

Question (1): What were some of the difficulties and causes of inaccuracy that you met when you were estimating scores for answers to your question?

Answer-Statements:

Among 15 scorers, only 4, constituting 27% of both groups, stated that there was no difficulty in estimating a score accurately, three were from the Mathematics Group and one was from the History Group. Eleven others, or 73% of both groups, mentioned certain difficulties they met in scoring papers.

The difficulties identified in scoring Mathematics papers were:

	Number of Respondents
(a) The explanation presented by students is, in many cases, not clear enough to make accurate estimation.	1
(b) Students' answers show that they misunderstood what is intended in the question.	1
(c) Several solutions of a problem are possible, certain solutions are easier to score than others.	1
(d) The checking of complicated arithmetical operations needed time and concentration.	1
(e) The question is general and subjective; the rating of an answer depends on personal judgement.	1

The difficulties identified in scoring History papers were:

	Number of Respondents
(a) The question is general, subjective, and require lengthy answers.	4
(b) Two forms of answers occured: the free essay form and the ennumerating-of-points form. The latter is easier to match with the scale; the former requires identifying the correct points as they occur in the essay.	1
(c) The question is the general knowledge type. The scorer is impressed by what of the student's general culture appears in the answer.	1

Question (2): To what extent did you subject your estimation of a score to the pre-established scale? Were you influenced by some elements in the answers of students other than those mentioned in the model answer?

Answer-Statements:

Although all the fifteen scorers stated that they adhered or tried to adhere to the pre-established scale, only five scorers (33%) who were from the Mathematics Group confirmed this adherence. The others (67%) expressed their deviations from the scale in different ways. These deviations can be classified into the following three categories:

(a) Deviations due to the ambiguity of the model answer and the scale on the mechanics of the answer, that is on features such as handwriting, organization, clarity of expression, sequence, and

any other feature aside from the knowledge part of the answer.

Each of 2 scorers from the Mathematics Group and 4 scorers from the History Group mentioned one or more of the above features that was considered either by giving it partial credit or by affecting the scorer's impression.

(b) Deviations due to deficiencies in the knowledge part of the model answer. This was admitted by one scorer from the History Group who stated that the subcommittee on the question recommended crediting those parts of the students' answers that look reasonable although they are not specified in the model answer, and although the total score points assigned to the question is already fully assigned to the model answer.

(c) Deviations that result from individual attitudes and personal impressions of scorers. These are illustrated by the following 4 answers made by 4 interviewees, the first answer by a scorer from the Mathematics Group; the others by three scorers from the History Group:

(1) The question was optional. Perhaps it was misleading because many students tried it but later they cancelled their answer to it. I agreed with my group to give part of the grade to cancelled correct answers.

(2) One external element affected my personal judgment; I kept in mind that a score of 8 (out of 20) is the minimum for passing. This had, in a way, approved all scores higher

than 8; almost all students passed the minimum of 8.

(3) Total impression and personal judgement affected the estimations to some extent.

(4) Sequence, logic, paragraphing, and organization are considered and have their effect on the scorer's impression.

Question (3): Did you notice from the answers you read that the question was ambiguous, misunderstood, or answered differently from what was intended?

Answer-Statements:

Each of 6 scorers from the Mathematics Group stated the question is clearly stated, there is no ambiguity in it.

All the 7 scorers of the History Group stated that the question is clearly stated -with some reservation added.

Criticisms of the questions were presented in different ways:

	Number of Respondents
(a) The wording of the question resulted in much guessing, more accurate words should have been used in stating the question. (Mathematics Group)	1
(b) The question was misunderstood by students. (Mathematics question)	1
(c) The question is subjective and requires lengthy answers. (History question)	3
(d) Although the question is general and subjective, the limits of the correct answer are very clear in the textbook.	4

Question (4): Do the scores on the answers you scored tend to be high, average, low, or varying?

Answer-Statements:

	Number of Respondents
The Mathematics Group:	
(a) Generally high (or tend to be high)	5
(b) Generally above average.[1]	1
(c) Vary	2
The History Group:	
(a) Tend to be above average.[1]	4
(b) Most answers got about the average.[1]	1
(c) Vary very much	1

It should be noted that the above results are only rough estimations of score tendency. Any inferences based on them have to be made with reservations. More evidence is needed to supplement the presence of a certain "tendency" in setting scores. However, it is worth observation that only three scorers stated that scores tend to vary; all others stated that the scores tend to be high, above average, or about the average (50% of maximum score), but no one stated that scores tend to be below average or low.[2] The scorers' conception of the average is worth considering. "It shows a predisposition to prejudge an average around which scores cluster. This is somewhere vaguely between standard scores (having fixed average) and absolute

---

[1] "Average" as used here means fifty percent of the maximum score on the question.

[2] The percentages of success in Mathematics and History are not available; the percentage of success in the whole Secondary Examination is over 70%.

achievement (indeterminate average). It has the disadvantages of both and the advantages of neither."[1]

Question (5): On what proportion of the papers does the "revisor" check the scores by going back to the answers and reading them again?

This question was addressed to the "revisor" working in conjunction with the scorer that was being interviewed. In a few cases the question was addressed to the scorer himself. However, since "revision" is supposed to have been performed on all the papers, the question seemed to present a challenge to the claim of "efficient performance" of the group of "revisors". This being the case there is reason to believe that the answers to the above question might not be accurate estimations of the proportion of papers being revised. The only conclusion that can be drawn is that a considerable proportion is not revised at all.

Among the Mathematics Group, five "revisors" reported that all the papers assigned to them are revised; two other scorers stated that 50% of the papers are revised at a much faster rate than original scoring. One scorer's reply was that since his question was an easy and straightforward one, he was entrusted to give the final score without being checked.

A group of "revisors" of History papers agreed among each other, when they were asked the question, that about 30% of the papers

---

[1] Professor F. Korf, Advisor of this thesis; Head of Office of Tests and Measurements, A.U.B., 1962-1963.

were revised.

The conclusions which can be drawn are:

(1) Over 50% of the Mathematics papers are usually revised, but definitely not all of the papers.

(2) About 30% of the History papers are usually revised.

(3) Revision of both Mathematics and History papers is done at a much faster rate than original scoring.

IV. Experiments on the Official Scoring of Standardized Student Answers

The purpose of these experiments is to investigate the consistency of the official scoring of Examination papers. More specifically it is to obtain an estimate of human error due to inconsistency among different scorers and to determine what effect this particular error has on the reliability of the Examination. Two types of "human errors" are differentiated; the one is due to inconsistency among different scorers (i.e., different grades assigned by different scorers to the same candidate's response) which is measured in these experiments, the other is due to the inconsistency of the same scorer (i.e., different grades assigned by the same scorer to identical responses graded at different times) and it is not measured in these experiments but the evidence provided by the interviews with scorers suggests its presence.

The method used in each experiment is to ascertain the variability of scores assigned to a paper by different scorers. In order that the results of these experiments have some validity in relation to what could be concluded about the Examinations, three conditions have been fulfilled: (1) The persons who participated in the experiments have been themselves scoring Examination-papers. (2) The standardized answers were real student answers in the same Examination whose papers were scored by persons participating in the experiments. (3) The experiments were conducted at the same

time, in the same halls, and using the same procedures of official scoring; that is they were performed under the same conditions as official scoring, and on regular scoring days.

It might be argued that the results of these experiments do not touch the accuracy of official scoring because in many cases one person scores all the answers to the same question; moreover, the system of "revision" assures maximum accuracy. In answer to this, it can be stated that:

First, in many other cases the answers to the same question are scored by more than one person.

Second, the system of "revision" in its present form is not very efficient and can be itself the cause of an error of the type investigated here since only a proportion of the papers is revised,[1] hence some scores are the result of one person's decision and others the result of agreement between two.

Third, by virtue of the fact that an answer lends itself to variation in the score assigned to it by different scorers, an error is introduced into every score even though the whole set of students' answers to a question is scored by one person, who is himself assumed to be free of any personal inconsistency from hour to hour and day to day.

The procedure: These experiments were performed on ten committees, five of which were scoring Secondary Examination papers and the other

---

[1] Supra, pp. 52-53.

five of which were scoring Preparatory Examination papers.

On each subject whose Examination papers were scored by one of these committees, a student answer was standardized in the following manner:

A question was selected from the question paper such that the answer to it is well known to all the members of the subject committee.

An answer booklet was drawn from the piles such that it contained an answer to the question selected.

The answer in the booklet was copied on stencil paper in such a way that the copy looked approximately the same as the original: the same handwriting, paragraphing, scratches, displaced words or figures, misspelled words, etc., all were imitated and transfered to the stencil. In addition, the question, the model answer and the scoring scale used by the subcommittee on the question, and an intro ductory statement at the beginning requesting each scorer to estimate independently a score for the "student answer", all were copied on stensil with the answer. Duplication was done to give a number of copies equal to the number of scorers in the corresponding committee. Copies of the standardized answer were distributed to the scorers after briefly explaining to them that the experiment is part of a research study on General Examinations and that their valuable participation lies in having each person score the "answer" given to him in the same way he scores Examination papers.

A modification of the experiment was tried with the committees of the Perparatory Examination. Since in the official scoring, a score is supposedly the result of agreement between two persons, a "corrector" and a "revisor", the members of each committee were asked, directly after collecting the papers of the first experiment, to score the same "answer" by having a "corrector" and the "revisor" agree among each other on a score. New copies of the standardized answer were used in the second experiment. However, the results of these "modified" experiments cannot be considered reliable because group discussion could not be avoided after the first experiment. It will be noted that the results of the second experiment are less inconsistent than the first. But because one of the influential factors, that is group discussion, was not controlled, the decrease in inconsistency cannot be totally attributed to the change in procedure.

The subjects on which standardized answers were prepared are:

A. Secondary Examination subjects: Mathematics, Physics, English Language, History, and Philosophy.

B. Preparatory Examination subjects: Arabic Language, English Language, Social Studies, Mathematics, and Sciences.

Statistical Considerations:

The measures of variability that are considered here are the range, variance, and the standard deviation. The formulas by which these statistics are computed are:

(1) The range = highest score - lowest score + 1

(2) Variance = $S^2 = \frac{i^2}{n}\left\{ \Sigma fx'^2 - \frac{(\Sigma fx')^2}{n} \right\}$

This formula is used for calculating the variance of grouped data, where:

i = interval between groups of scores (between class-marks).
f = frequency of scores within one interval (within class-boundaries).
x' = deviation of a score from the arbitrary reference point of grouped scores.
n = number of scores.
$\Sigma fx'^2$ = sum of squares of deviations from the arbitrary reference point.
$\Sigma fx'$ = sum of deviations of scores from the arbitrary reference point.

(3) Standard deviation = $S = \sqrt{\text{Variance}}$

The mean of each sample of scores is calculated by the formula,

(4) $\bar{X} = m' + \frac{\Sigma fx'}{n}(i)$

where, $\bar{X}$ = the mean, m' = arbitrary mean (arbitrary reference point of grouped scores). The other notations have the same meanings as above.

The Results of the Experiments

A. The General Secondary Examination:

(1) English Language:

Scores assigned by 30 scorers to 30 identical copies of a single student's answer to a 10-point question on English Prose: (Numbers in parentheses show number of scorers assigning that score)

8(2), 6(2), 5(6), 4(10), 3(8), 2(2).

Range = 7 , $\bar{X}$ = 4.20 , S^2 = 2.02 , S = 1.42

(2) History:

Scores assigned by 45 scorers to 45 identical copies of a single student's answer to a 20-point question on Arab History: (Numbers in parentheses show number of scorers assigning that score)

15(4), 14.5, 14(7), 13(6), 12.5(2), 12(6), 11.5(2), 11(5), 10.5(2), 10(5), 9.5(1), 9(3), 7(1)

Range = 9 , $\bar{X}$ = 12.00 , S^2 = 3.68 , S = 1.92

(3) Social Studies:

Scores assigned by 16 scorers to 16 identical copies of a single student's answer to a 14-point question on Philosophy: (Numbers in parentheses show number of scorers assigning that score)

12(1), 11(2), 10(4), 9(6), 8(1), 7(1), 5(1).

Range = 8 , $\bar{X}$ = 9.25 , S^2 = 2.56 , S = 1.60

(4) Mathematics:

Scores assigned by 24 scorers to 24 identical copies of a single student's answer to a 25-point question on Algebra: (Numbers in parentheses show number of scorers assigning that score.)

23(1), 22(2), 20(2), 19(4), 18(6), 17(5), 16(2), 15(1), 13(1).

Range = 11 , $\bar{X}$ = 18.17 , S^2 = 4.80 , S = 2.19

(5) Physics:

Scores assigned by 12 scorers to 12 identical copies of a single student's answer to a 14-point question on Magnetic Measurement: (Numbers in parentheses show number of scorers assigning that score.)

9(1), 8(5), 7(1), 6(4), 5(1).

Range = 5 , $\bar{X}$ = 7.08 , S^2 = 1.41 , S = 1.19

B. The General Preparatory Examination:

(1) Arabic Language:

Scores assigned by 23 scorers to 23 identical copies of a single student's answer to a 15-point question on Arabic Literature: (Numbers in parentheses show number of scorers assigning that score.)

14(1), 13(3), 12(5), 11(9), 10(5).

Range = 5 , $\bar{X}$ = 11.4 , S^2 = 1.20 , S = 1.09

(Scores assigned by 11 subcommittees, each consisting of a "corrector" and a "revisor", to the same answer on Arabic Literature: 12(6), 11(5).

Range = 2 , $\bar{X}$ = 11.54 , S^2 = 0.21 , S = 0.46)

(2) English Language:

Scores assigned by 24 scorers to 24 identical copies of a single student's answer to a 10-point question on Translation: (Numbers in parentheses show number of scorers assigning that score.)

10(1), 9(3), 8(5), 7.5(2), 7(10), 6(3).

Range = 5 , $\bar{X}$ = 7.50 , S^2 = 0.98 , S = 0.99

(Scores assigned by 12 subcommittees to the same answer on English Translation: 8.5(1), 8(4), 7.5(3), 7(4).

Range = 2.5 , $\bar{X}$ = 7.58 , S^2 = 0.32 , S = 0.57)

(3) Social Studies:

Scores assigned by 23 scorers to 23 identical copies of a single student's answer to a 16.5-point question National Education: (numbers in parentheses show number of scorers assigning that score.)

16(2), 15.5(1), 15(4), 14.5(2), 14(1), 13(1), 12.5(1), 12(2), 11(2), 10.5(1), 10(1), 9.5(2), 8(2), 5(1).

Range = 12 , $\bar{X}$ = 12.28 , S^2 = 8.58 , S = 2.45

(Scores assigned by 11 subcommittees to the same answer: 15.5(2), 15(2), 14(1), 12.5(1), 12(2), 10(1), 9(1), 8.5(1).

Range = 8 , $\bar{X}$ = 12.73 , S^2 = 6.00 , S = 2.45)

(4) Sciences:

Scores assigned by 16 scorers to 16 identical copies of a sigle student's answer to a 10-point question on Chemistry: (Numbers in parentheses show number of scorers assigning that score.)

10(1), 9(2), 8(3), 7.5(2), 7(6), 6(2).

Range = 5 , $\bar{X}$ = 7.56 , S^2 = 1.09 , S = 1.04

(scores assigned by 8 subcommittees to the same answer on Chemistry: 8(1), 7.5(1), 7(5), 6.5(1).

Range = 2.5 , $\bar{X}$ = 7.06 , S^2 = 0.09 , S = 0.30)

(5) Mathematics:

Scores assigned by 18 scorers to 18 identical copies of a single student's answer to a 14-point question on Arithmatic:

(Numbers in parentheses show number of scorers assigning that score.)

11(3), 10(10), 10.5(1), 9(1), 7(1), 4(1), 3(1).

Range = 9 , $\bar{X}$ = 9.22 , S^2 = 4.84 , S = 2.20

(Scores assigned by 9 subcommittees to the same answer on Arithmatic: 11(2), 10(6), 7(1).

Range = 5 , $\bar{X}$ = 9.89 , S^2 = 1.21 , S = 1.10)

V. Analysis of Random Samples of Examination Grades

To determine the extent to which the reliability of Examinations is affected by such estimates of human error (due to inconsistency among scorers) as those obtained in the above experiments, the actual distribution of scores on the Examinations is needed. Because of the difficulty involved in getting the whole distribution, a random sample of scores on each of the subjects represented in the experiments was obtained. The variability of scores in each sample is computed and the values obtained are considered as estimates of the variability of the whole distribution. The way in which the reliability, the score variability, and the error variability are interrelated will be discussed in section VI.

The Method of Sampling:

A. Secondary Examination Grades: The Assistant Secretary of the Examinations reported that the lists of grades had the names of students arranged randomly for the purpose of assigning random secret numbers. Therefore, starting with the first name on the list, every thirtieth name in the scientific section, and sixtieth name in the literary section was chosen. A sample of 35 grades on each of the two science subjects, and of 51 grades on each of three literary subjects was collected. These numbers correspond to about 1100 participants in the scientific section Examination,

and to about 3100 participants in the literary section Examination.

B. Preparatory Examination Grades: The list of grades had the names of students arranged according to their ranks in the Examination. At the same time, the names were given serial numbers from 1 to 2902.[1] To obtain random samples of grades, a number was selected randomly from the column of ranks; this number was read in the column of serial numbers, and the grades on five subjects belonging to it and to every fiftieth number succeeding and preceeding it in the serial order were taken. A sample of 58 grades on each of five Examination topics was thus collected.

The measures of variability of scores in every sample will be computed in the same way as illustrated in section IV pp. 57-58.

A. Samples of Secondary Examination Grades:

(1) English Language:

28	95	190	76	133	216	166	127	147
106	143	161	50	174	196	94	65	36
179	144	135	120	164	238	170	99	196
175	154	114	154	142	175	116	188	
89	197	127	97	160	79	46	63	
80	157	65	131	158	132	128	125	

Range = 211 , $\bar{X}$ = 130.3 , S^2 = 1954 , S = 44.2

(2) Arab History:

25	48	33	61	60	53	33	49	77	56	44	61	70
74	61	55	56	54	67	60	60	50	65	64	70	51
47	61	79	52	50	54	64	56	40	71	67	69	82
57	53	72	49	34	59	45	74	58	66	40	66	

Range = 58 , $\bar{X}$ = 57.39 , S^2 = 151.3 , S = 12.3

---

[1] This is the number of participants in the Preparatory Exam. from the Balka District only.

(3) Social Studies:

25	59	86	59	71	53	59	28	30	67	72
69	48	59	62	58	49	91	60	61	56	
43	41	73	63	45	40	95	69	31	73	
72	44	47	40	62	60	51	65	29	60	
34	80	76	49	33	81	6	41	49	41	

Range = 90 , $\bar{X}$ = 50.23 , S^2 = 331.2 , S = 18.2

(4) Mathematics:

146	194	243	161	180	253	122	220	87
225	41	141	254	231	132	147	266	86
91	176	246	128	162	185	228	174	31
130	223	245	240	221	129	204	231	

Range = 236 , $\bar{X}$ = 177.93 , S^2 = 3919 , S = 62.6

(5) Physics:

52	51	70	67	73	80	45	57	37
75	4	40	72	60	50	68	96	43
40	59	71	56	51	76	68	56	4
41	65	79	77	71	43	66	85	

Range = 93 , $\bar{X}$ = 59.0 , S^2 = 400 , S = 20.0

B. Samples of Preparatory Examination Grades:

(1) Arabic Language:

82	74	71	64	74	71	63	59	50	54
77	75	75	62	64	70	64	63	61	42
87	76	76	62	69	65	58	57	69	55
85	65	74	74	64	53	60	60	54	43
80	79	60	61	64	71	66	56	61	
69	72	77	62	76	69	60	60	50	

Range = 46 , $\bar{X}$ = 65.76 , S^2 = 92.16 , S = 9.6

(2) English Language:

83	69	69	59	29	43	62	45	38	50
78	83	67	69	56	43	50	33	35	24
72	84	71	78	67	55	36	47	36	20
60	81	76	63	50	5	19	42	10	48
71	73	66	74	59	53	30	19	36	
84	36	67	51	73	45	44	47	52	

Range = 80 , $\bar{X}$ = 53.03 , S^2 = 374.5 , S = 19.35

(3) Social Studies:

84	62	72	56	74	71	64	60	58	56
65	68	67	65	69	68	56	63	40	43
73	70	67	67	66	68	62	38	48	36
72	67	44	58	55	44	51	60	51	28
75	65	63	56	52	44	47	55	55	
53	73	50	64	60	45	59	53	36	

Range = 57 , $\bar{X}$ = 58.55 , S^2 = 130.0 , S = 11.4

(4) Sciences:

88	67	58	67	25	21	50	45	42	17
84	54	61	41	67	28	45	40	33	26
80	75	47	41	50	38	54	36	27	10
77	64	51	55	29	66	40	34	40	15
69	76	61	58	50	59	37	23	35	
70	61	70	38	41	38	15	40	20	

Range = 79 , $\bar{X}$ = 47.86 , S^2 = 352.7 , S = 18.78

(5) Mathematics:

90	82	75	71	16	50	7	4	4	24
86	72	52	68	62	30	21	4	23	3
72	56	38	50	50	5	39	14	4	26
44	76	37	50	41	55	36	29	7	11
77	43	29	16	37	19	50	18	42	
98	27	60	67	31	68	37	25	8	

Range = 96 , $\bar{X}$ = 40.36 , S^2 = 672 , S = 25. 92

VI. Interpretation of Results

The results of investigations described in this chapter provide two types of evidence:

A. the presence of elements that tend to reduce the validity of the Examinations, and

B. the presence of elements that tend to reduce the reliability of the Examinations.

A. The Validity of the Examinations:

The type of validity on which some evidence could be collected is "content validity". Another type of validity on which no studies have been made, but can be of particular importance, is "predictive validity". In order that public examinations can be effectively used for selection purposes, they must possess predictive validity, that is, they must successfully predict future achievement in areas of study to which students are selected. "Content validity refers to how well an examination represents the content of a course of study. The content validity of an essay examination can be judged by: (1) the extent to which thoughtful planning has been put in setting the questions, (2) how well the questions chosen match with course objectives, (3) the scope and quality of text-books and other sources used in setting the questions, (4) how much sampling of the course content was brought in the questions, and (5) the choice of

specialized and experienced persons for setting the questions.

The elements revealed by the investigations that would contribute to the validity of the Examinations are:

(1) The questions were prepared by subject-specialists who had experience in teaching and in school-inspection.

(2) The questions were confined to the scope of the prescribed syllabus. The prescribed textbooks, which are no more than the syllabus in detailed form, were used by most writers of questions.

The elements revealed by the investigations that tend to reduce content validity are:

(1) There was little thoughtful planning of the questions according to objectives defined and analyzed prior to the setting of the questions. There was organization of the work among the members of a group appointed to write a set of questions, but this organization consisted, in most cases, of mere partitioning the subject-content among the members. This is likely to produce unfavorable elements of heterogeneity within the same set of questions, which might result in overlapping and in some cases conflicting mental processes anticipated from students upon answering the questions. The point to be made here is that poor planning is not likely to produce what might be considered an appropriate representation of the course content.

(2) The use of sources other than textbooks was very limited. Other sources of the material could have been very enlightening on certain aspects that can better lend themselves to testing what might be considered as basic abilities and basic concepts.

(3) Class instruction was ignored to a very large extent. Confining the questions to bare topics outlined in the syllabus can be no more than partial testing. Even the reliance on textbooks is not very reliable because most of the textbooks available present the material in a formal, reportorial style. Hence an essential part of the course content is normally supplied by the teachers. That class instruction should be considered as a part of the course content is "officially" supplemented by the fact that the syllabus includes, in addition to the outline of the course, a set of directions for class-room instruction and a set of behavioral objectives for the course. The investigations have shown that what teachers actually teach in their classes was not very clear in the minds of the writers of examination-questions at the time of the writing.

(4) The examination-questions included a relatively narrow sampling of subject content. Rough comparisons of sets of Examination-questions with the prescribed course content as outlined in the syllabus or the textbook shows that less than one-fourth of the content is represented in the questions of the Secondary Examination and less than one-sixth of the content is represented in the questions of the Preparatory Examination.

(5) The criteria by which some question-writers selected the material for their questions are not valid in that they do not correspond to the actual purposes of General Examinations. For example, "the actual level of attainment" as a criterion maintained by some question-writers is not in accord with "standard level of achievement" as a general purpose of the Examinations. If the actual level of attainment of students in subject X is low, the examination on subject X designed according to this criterion would measure a lower level than the prescribed standard, or it might measur a different thing from what is actually prescribed.

On the other hand, a criterion such as "the essential points in the text" or "basic concepts and ideas" can be regarded as valid if "essential points" or "basic concepts and ideas" can be recognized as the same things by every one, or if the group of specialists on the subject came to an agreement about them. That these criteria are not recognizable as the same things by every one is evident from the contrast between "essential points" and "basic concepts and ideas". That no agreement among specialists was worked out is evidenced by some question-writers stating that they had been the main persons who did the work of question-writing, and that their colleagues did not participate systematically.

All this confirms the conclusion that General Examinations, with the type of essay questions set for them, are testing only a restricted proportion of course content, the representativeness of which suggests low content validity.

B. The Reliability of the Examinations:

Defined in general terms, the reliability of an examination refers to how consistently the examination measures whatever it does measure. The reliability can be estimated quantitatively by ascertaining the accuracy of measurements, or scores, on a test or examination, and a "reliability coefficient" is arrived at. Several methods are used to calculate reliability coefficients. The methods used in this study will be considered in the course of discussion.

In an essay examination, the main sources of unreliability are: (1) The subjectivity of the questions, and (2) The subjectivity of scoring. The investigations on General Examinations in Jordan described in this chapter have yielded evidence on both:

(1) Subjectivity of the Questions:

(a) Certain objectives exhibited in some questions and testified to by some of the question-writers make the questions highly subjective. Examples of such objectives are mentioned on page 39. Examples of subjective questions are given on pages 40-41.

(b) Many question-writers, especially those on languages and social studies, expressed their tendency to write general and subjective questions in which they expect "the student to express himself in his own way", "to show quality of ideas", "to show ability to present a good argument", etc.

The tendency of question-writers to write general and subjective questions is explicitly expressed in the way they phrased their

questions. It was shown that about 70% of the questions on the General Secondary Examination (the Literary Section) and about 45% of the questions on the General Preparatory Examination use phrases such as "Talk about"; "Write about"; "Explain in brief, or detail"; "What is your opinion"; and a few other phrases, all of which tend to make questions open, indefinite, and subjective.[1]

(c) That some question-writers did not prepare model answers can be interpreted that a chance of reconsidering the questions and making them, perhaps, more definite has escaped those writers. Moreover, some question-writers admitted that they did not want to impose their "personal opinions" on the members of the scoring committees as to what should be the correct answer. Such writers are self-confessedly aware of the subjectivity of their questions.

(d) There was recognition on the parts of scorers, including those of mathematics papers, that the questions contained elements of subjectivity. Some of the statements expressed by scorers describing the subjectivity of the questions were mentioned on pages 47-48.

(2) Subjectivity of Scoring:

(a) The majority of scorers interviewed mentioned that certain elements and characteristics such as handwriting, organization, clarity of expression, exposition of ideas, etc., which were not specified in the model answer or the scoring scale, have influenced their scoring. In some cases these characteristics were given

---

[1] Supra, p. 42.

partial credit, in others they affected the total impression. That different characteristics were considered by different scorers can be interpreted as one cause for the variation among different scorers. On the other hand, the inconsistency of the same scorer is shown by the fact that any of these characteristics cannot be sharply classified or differentiated into categories that can be easily rated. What exactly are the types of handwriting, of organization, of exposition of ideas, etc. that a scorer comes across when he reads paper after paper, and how is each type rated? It is doubtful that any scorer will go into the trouble of making such fine classifications and ratings inorder that he will be consistent throughout. Furthermore, the inconsistency of the same scorer is grossly magnified when certain attitudes of scorers are considered: the attitude of a scorer who complains of checking tedious arithmetical operations, of another who says that the answers to his question tend to be very lengthy, of a third who is impressed by the general culture of the student, of a fourth who deliberately bears in mind the minimum points for passing the question so that almost all the papers he scores pass, etc. The last example is of particular significance because it explicitly means that the scorer was not nearly as concerned about a score's accuracy as he was about its being above the passing grade.

All this will strongly supplement the conclusion that any score resulting from the examinations will contain a significant error, due in part to the inconsistency of different scorers and in

part to human inconsistencies within the same scorer.

(b) The system of "revision" in its present form is not very effective and can be itself the cause of amplifying errors resulting from variations among different scorers, since only a proportion of the papers is revised, hence, on the same set of answers to a certain question, some scores are the result of one person's decision and others the result of agreement between two.

Another deficiency of revision is that the "revisor" does not read the answer as carefully as the "corrector" does. Moreover, he has the chance of noting the corrector's score recorded on the booklet, which is quite likely to bias his judgement. Many revisors follow the procedure of selecting for revision only the papers carrying very low or very high corrector's scores. Some try to be sympathetic by looking for points in the low-scored answer that can be credited, and thus raise the low score, while others try to be more strict in checking the high-scoring answer. The attitude of sympathy can be interpreted to be one of the causes why most scorers of Mathematics and History papers reported that scores tend to be "generally high" or "above average". This conclusion cannot be generalized without having more valid evidence, because it can be claimed that other causes can lead to high scores. However, coupled with other evidence mentioned above, this result indicates that the hypothesis that some scorers are consistently lenient and others consistently severe can probably be verified by valid procedures.

(c) The results of the experiments on the official scoring of Examination-papers supplemented the conclusion that any score on the Examinations contains a significant error due at least in part to inconsistency of scoring.

The way in which this error affects the reliability of the Examinations is illustrated in the following treatment adapted from Downie and Heath:[1]

A score is considered as being made up of two parts, a true score component and an error score component. This can be expressed as:

$$X = X_t + X_e \qquad (1)$$

where:

X = any raw score
X_t = true score component
X_e = error score component

The variance of a test can be treated similarly as being made up of the variance associated with true scores (true variance) and the variance associated with error scores (error variance). This can be expressed as:

$$S^2 = S_t^2 + S_e^2 \qquad (2)$$

where:

S^2 = test variance
S_e^2 = error variance
S_t^2 = true score variance

---

[1] Adapted from N. M. Downie and R. W. Heath, Basic Statistical Methods, pp. 190-197. Harper & Brothers, Publishers, New York, 1959.

By dividing equation (2) through by a constant S^2, and transposing:

$$\frac{S_t^2}{S^2} = 1 - \frac{S_e^2}{S^2} \qquad (3)$$

As reliability is defined as that part of the variance which is true variance, then:

$$r_{tt} = \frac{S_t^2}{S^2} \qquad (4)$$

where, r_{tt} = reliability of the test.

or,

$$r_{tt} = 1 - \frac{S_e^2}{S^2} \qquad (5)$$

Equation (5) can be written in another form:

$$S_e = S\sqrt{1 - r_{tt}} \qquad (6)$$

S_e in equations (5) and (6) is the standard deviation of a sample of scores of an individual about his true score. It is called the standard error of measurement.

Now if the error due to inconsistency among different scorers that has been determined experimentally were the only error introduced in a score, then it would be considered as an estimate of the standard error of measurement S_e, and since the standard deviation of Examination scores is estimated by the standard deviation of the random samples of Examination scores, the Examination-reliability would be determined by using equation (5) above. But there are two other types of error that have to be taken into account. One of these errors has already been mentioned

as human error due to the inconsistency of the same scorer. The evidence derived from interviews with scorers suggest that this error can be of considerable value. The other error is that which exists in a test or examination after eliminating human error entirely. This error is due to variations within the examination content and within the examinee. It varies from one test to another and has minimum value in an objective test of very high reliability. Usually it is this error which is denoted by S_e in equation (5). It can be noted from this equation that a large value of the standard error of measurement corresponds to a low reliability, and visa versa.

Now in an essay-examination, where all three types of error are encountered, the standard error of measurement S_e must include them all.[1] Hence, if

S_{e1}^2 = error variance due to variations in the test and in the individual,

S_{e2}^2 = error variance due to variations within the same scorer,

and S_{e3}^2 = error variance due to variations among different scorers,

then,

$$S_e^2 = S_{e1}^2 + S_{e2}^2 + S_{e3}^2 \qquad (7)$$

Equation (2) above becomes:

$$S^2 = S_t^2 + (S_{e1}^2 + S_{e2}^2 + S_{e3}^2) \qquad (8)$$

---

[1] Professor F. Korf, Advisor of this thesis, Head of Office of Tests and Measurements, A.U.B., 1962-1963.

In the experiments on official scoring, only S_{e3} has been numerically determined. The other two errors, S_{e1} and S_{e2} are numerically unknown but descriptive evidence on their presence in significant values has been described earlier in this section. The size of these two errors can be estimated by ascribing to the Examinations a reasonable value of reliability. This value is suggested on the basis of findings of other more elaborate studies on essay examinations.

The earliest of these studies, made by Starch and Elliot in 1913,[1] used a procedure similar to the one used in the aforementioned experiments, that is, the unreliability of an essay-examination was shown merely by the variability of scores assigned to a paper by different scorers.

Later studies used a correlation method. Two forms of an essay examination were given to the same students and marked by experienced examiners. The average correlation between two sets of scores assigned by the same examiner to the two forms was taken as the coefficient of reliability of the examination.[2] On the other hand, the average correlation between two sets of scores assigned to the same examination by two examiners was taken as the coefficient of reliability of marking the examination. McGreger and Rush used

---

[1] Walter S. Monroe, Encyclopedia of Educational Research, p. 408, Editor: Walter S. Monroe, The Macmillan Co. New York, 1952.

[2] W. S. Monroe, J.C. DeVoss, and F. J. Kelly, Educational Tests and Measurements, p. 470, Houghton Mifflin Co., New York, 1924.

this procedure in studying eigth-grade examinations in sixteen subjects from 952 pupils in eleven states.[1] Each paper in the two sets of examinations was marked independently by two experienced teachers. An average correlation of .62 was obtained for marking, and of .43 for the examination.[2] A similar study by Gorden yielded a marking reliability of .72 and an examination reliability of .42.[2] Another study conducted at the University of Chicago High School showed a marking reliability of .94 on Form A and .84 on Form B, but examination reliability (correlation between Form A and Form B) was only .60.[2]

These studies in which recommended procedures of marking were used show that the examination-reliability is considerably lower than marking reliability.

Monroe and Saunders reported 66 coefficients of essay examination reliabilities with a median of .65.[3] Rush summarized 285 coefficients of essay examination reliabilities including those obtained by Monroe an Saunders. Nine fell in the interval .93 to .97 and nearly as many were negative; the median was .59.[3]

Monroe thinks it is appropriate "to interpret the total evidence as indicating that, for a wisely prepared one hour [essay] examination, an examination reliability of .65 to .70 may be anticipated when recommended procedures are employed in determining the marks assigned."[3]

---

[1] C. C. Ross, Measurement in Today's Schools, p. 159, New York, Prentice-Hall, Inc., 2nd edition, 1953.

[2] Ibid, pp. 159-160.

[3] Encyclopedia of Educational Research, loc. cit.

It should be noted that in all these studies where "recommended procedures of marking were employed", the attempt has been to determine the "content" reliability of an essay examination, which is equivalent to the reliability of a keyed objective test that has a perfect marking reliability of 1.0. It also means that for accurately determining the content reliability of an essay examination, a high marking reliability must be assured, since as it has been already mentioned, marking reliability is an upper bound for the examination reliability.

This leads to the conclusion that if General Examinations in Jordan were "wisely prepared" and the marking of papers highly reliable, then an examination-reliability of .65 to .75 would be expected—adopting Monroe's interpretation of previous findings. With the evidence brought forth in this chapter about the subjectivity of the questions and the subjectivity of scoring, a "net" examination-reliability lower than .65 should be expected.

Taking an average examination-reliability of .65 as an unprejudiced value of the reliability of General Examinations in Jordan, the other part of human error (S_{e2}) and the objective error (S_{e1}) can now be computed. The results of computations are shown in Table 3. The method of computation is illustrated in the following example in which the results of investigation on the mathematics test of the Preparatory Examination are considered:

(1) From the experiments, the standard deviation of scores assigned by different scorers to the standardized answer is considered as being an estimate of the human error due to variations among scorers.

Hence, for a 14-point Mathematics question:

Standard error = $\underline{2.2}$, error variance = $(2.2)^2 = \underline{4.84}$

Assuming that all the questions in the Mathematics paper are parallel, the error variance of the entire Mathematics test will be equal to the error variance of the 14-point question (considered as a 14-point test) multiplied by the ratio of the total points on the whole test to the total points on the question.[1]

Hence for the 100-point Mathematics test:

Error variance = $S_{e3}^2 = 4.84 \times \frac{100}{14} = \underline{34.6}$, $S_{e3} = \underline{5.9}$

(2) From the analysis of random samples of scores,

Standard deviation of Mathematics test = S = 25.92

Mathematics test variance = $S^2 = 672$

By substituting $r_{tt} = .65$ and $S^2 = 672$ in the equation

$$r_{tt} = 1 - \frac{S_e^2}{S^2}$$

we obtain, $S_e^2 = \underline{235} = S_{e1}^2 + S_{e2}^2 + S_{e3}^2$ = total error variance

$S_e = \underline{15.3}$

But, $S_{e3}^2 = 34.6 \simeq 35$

Hence $S_{e1}^2 + S_{e2}^2 = \underline{200}$

---

[1] H. Gulliksen, Theory of Mental Tests, p. 73, New York, John Wiley & Sons, Inc., 1950.

(3) The value of S_{el} is computed on the assumption that the essay test, if freed from human error, would behave as an objective test with a reliability as high as .80, which is flattering to an essay test, and with a distribution of scores approaching that of CEEB standard achievement tests, in which the standard deviation is about one-sixth the maximum range.[1] Hence for the 100-point Mathematics test, the standard deviation of scores freed from human error is about 16; score variance is, therefore, 256.

By substituting $r_{tt} = .80$ and $S^2 = 256$ in the equation

$$r_{tt} = 1 - \frac{S_e^2}{S^2}$$

we obtain: $S_e^2 = 51.2 \simeq \underline{51} = S_{el}^2$ for the essay test

and, $S_{el} = 7.15$

Hence? $S_{e2}^2 = 200 - 51 = \underline{149}$

and, $S_{e2} = 12.4$

(4) Summary of results:

S_e	S_{el}	S_{e2}	S_{e3}	S_e^2	S_{el}^2	S_{e2}^2	S_{e3}^2
15.3	7.15	12.4	5.9	235	51	149	35

Note: Table 3 gives the error variance of scores on each test, the standard error can be computed by taking the square root of error variance.

---

[1] This approach is suggested by Professor F. Korf, the advisor of this thesis and the Head of the Office of Tests and Measurements , A.U.B., 1962-1963. CEEB = College Entrance Examination Board.

Table 3

Components of Error Variance of Scores

on General Examinations in Jordan

(for an estimated examination reliability of .65)

	Error Variance			
	Total	Equiv. Obj. Test[1]	Human	
	S_e^2	S_{e1}^2	S_{e2}^2	S_{e3}^2
Sec. Examination				
English Language	682	500	121	61
History	53	51 ?	?	18
Social Studies	115	51	46	18
Mathematics	1370	500	812	58
Physics	140	51	79	10
Prep. Examination				
Arabic Language	32	51 ?	?	8
English Language	130	51	70	10
Social Studies	46	51 ?	?	52
Sciences	123	51	62	11
Mathematics	235	51	149	35

[1] S_{e1} is computed for an equivalent objective test of .80 reliability.

Notes on Table 3:

This table is computed on the assumption that the "content" reliability of each test could be as high as .80 and that this value was reduced to .65 by the introduction of human error. The value .65 is taken as the average reliability of all the tests included in either the Preparatory or the Secondary Examination. Yet there are grounds to reason that any of the constituent tests would hardly exceed an interval of reliability of more than .60 to .70. Table 3 shows clearly that the tests on Arabic and Social Studies in the Preparatory Examination and the test on History in the Secondary Examination should have a reliability less than .65, since 35% of raw score variance which was taken as total error variance is much less than the sum of the equivalent objective test error variance for .80 reliability and the human error variance verified experimentally. The unreliability of these tests is further shown by the comparatively high values of the human error variance S^2_{e3}, and by the small range of raw Examination scores compared to total test points. The ratio of the range to total points in all of the above-mentioned three tests is less than .6 .[1] This indicates that a good portion of each of these three tests was not discriminating at all. The region of unreliability can be identified by noting the position of the mean score on the test distribution on pages 64-66. Except for one test in the Preparatory Examination

[1] Refer to pages 64 to 66 for the range of scores on each subject.

and two tests in the Secondary Examinations that have a ratio of range to total test points more than .9, all other tests have this ratio less than .8, which means that these latter tests have considerably large areas of indiscrimination, hence unreliability.

One further observation on Table 3 is that the values of the human error variance S^2_{e2} in most cases look unrealistically high. It is to be noted that the value of S^2_{e1} was obtained by subtracting the sum of S^2_{e1} and S^2_{e3} from the total error variance S^2_e. Hence S^2_{e1} could be of larger value due to a "content" reliability less than the assumed .80. A larger value of S^2_{e1} will readjust the value of S^2_{e2} to a lower one. On the other hand, S^2_{e3} for a whole test was computed on the assumption that all the questions in a paper are parallel. This assumption can be questioned in several cases: In the English paper of the Preparatory Examination, all the questions were considered parallel to a straightforward "translation" question, when the paper contained questions on composition, reading, etc., that are far more subjective than the translation question, and that would have produced a larger value of S^2_{e3} if, say, a question on "prose" were included in the experiment. A similar situation occurs with other tests. However, there is some reason to believe that if the experiments were more comprehensive, the value of S_{e3} would be larger than the one already obtained. A larger value of S^2_{e3} would readjust the estimated value of S^2_{e2} to a lower one.

The objection that might be raised would be against designating such a low value as .65 to the reliability of the Examination.

It can be argued that it is more reasonable to assume a high value of reliability which means that the total error variance is smaller and hence the constituent parts of error variance would look more reasonable.

It is to be admitted that the actual reliability of the Examinations is still unknown. The value .65 was taken as a reasonable value proposed on the basis of findings by previous researches and confirmed by the evidence resulting from the investigations on the Examinations. The reasonableness of the value .65 can be seen when the same procedure of computation followed previously for estimating the values in Table 3 is carried out by considering other values of examination-reliability. The results of such computations are shown in Tables 4 and 5. In these tables it can be noted that for every test a point is reached where a certain proportion of the test variance is not sufficient to cover the sum of the error variance corresponding to an objective test of .80 reliability and the human error variance (S^2_{e3}) identified experimentally, thus ignoring the other human error S_{e2} totally. For example, in Table 5, a reliability of .40 taken for the Arabic test means that 60% of the test variance or 55 (= 92X60/100) is not sufficient to cover the error variance of 51 taken when the Arabic test is treated as an objective test of .80 reliability plus the human error variance of 8 identified experimentally. Hence the reliability of the Arabic test is most likely less than .40. The same reasoning is carried out to other tests, by which one arrives

Table 4

Components of Error Variance

of Scores on General Secondary Examination

for Different Values of Examination- Reliability

Subject	r_{tt}	Error Variance				
		Total	Equiv. Obj. Test[1]	Human		
		S_e^2	S_{e1}^2	S_{e2}^2	S_{e3}^2	
English Language	.50	975	500	414	61	$r_{tt} < .70$
	.60	760	500	199	61	
	.70	585	500	24 ?	61	
	.75	476	500	?	61	
History	.40	91	51	22	18	$r_{tt} < .50$
	.50	75	51	6 ?	18	
	.55	68	51	?	18	
Social Studies	.60	132	51	63	18	$r_{tt} < .75$
	.70	99	51	30	18	
	.75	82	51	13 ?	18	
	.80	66	51	?	18	
Mathe-matics	.70	1160	500	618	58	$r_{tt} < .85$
	.80	784	500	226	58	
	.85	587	500	29 ?	58	
	.90	392	500	?	58	
	.70	120	51	59	10	$r_{tt} < .85$
	.80	80	51	19	10	
	.85	60	51	?	10	
						Average $r_{tt} < .73$

[1] (S_{e1}) is computed for an equivalent objective test of .80 reliability.

Table 5

Components of Error Variance

of Scores on General Preparatory Examination

for Different Values of Examination-Reliability

Subject	r_{tt}	Error Variance				
		Total	Equiv. Obj. Test[1]	Human		
		S_e^2	S_{el}^2	S_{e2}^2	S_{e3}^2	
Arabic Language	.30	64	51	5	8	$r_{tt} < .35$
	.35	60	51	1 ?	8	
	:40	55	51	?	8	
English Language	.70	111	51	50	10	$r_{tt} < .80$
	.80	74	51	13	10	
	.85	56	51	?	10	
Social Studies	.20	104	51	1	52	$r_{tt} < .20$
	.30	91	51	?	52	
Sciences	.60	140	51	78	11	$r_{tt} < .80$
	.70	105	51	43	11	
	.80	70	51	9	11	
	.85	53	51	?	11	
Mathematics	.70	202	51	116	35	$r_{tt} < .85$
	.80	134	51	48	35	
	.85	100	51	14	35	
	.90	67	51	?	35	
						Average $r_{tt} < .60$

[1] S_{el} is computed for an equivalent objective test of .80 reliability.

at the conclusion that an average reliability of less than .60 is expected for the Preparatory Examination and an average reliability of less than .75 is expected for the Secondary Examination. It must be remembered, however, that in making this conclusion there was no mention of findings by previous researchers nor of evidence resulting from the investigation on the Examinations. Therefore, Tables 4 and 5 give exaggerated values if relied upon for determining the reliabilities of the different tests.

Lastly it is worth mentioning how the experimental values of S_{e3} alone have affected the reliability of the Examinations. Since reliability is defined as that part of test variance which is true variance, any increase in the ratio of error variance to the test variance corresponds to an equal decrease in test-reliability, and visa versa.

By taking the ratio of the error variance (S^2_{e3}) to the corresponding test variance and expressing the result as a percentage, the following results are obtained:

A. The Secondary Examination:

English	History	Social Studies	Mathematics	Physics
3.1%	12.2%	5.6%	1.5%	2.5%

Average = 5% approximately

This means that the introduction of human error due to inconsistency among scorers into the examination has, on the average, decreased a pre-supposed reliability by .05, and it means, too, that the elimination of this error from the examination would, on the

average, have increased the reliability by .05.

B. The Preparatory Examination:

Arabic	English	Social Studies	Sciences	Mathematics
8.7%	2.7%	40%	3.1%	5.2%

Average = 12% approximately

(Excluding Social Studies, the average = 5% approximately)

Summary of Conclusions

Investigations on the General Examinations in Jordan have shown that:

(1) The way in which questions are set suffer from weaknesses that tend to reduce the content validity of the Examinations.

(2) The questions themselves include a significant degree of subjectivity and indefiniteness.

(3) The scoring of papers involves a significant degree of subjectivity which allows for two types of human error, the one due to inconsistency of different scorers and the other due to inconsistencies within the same scorer. A third and unavoidable error persists in any case when the Examinations are regarded as tests free from human error.

(4) The subjectivity of the questions and the subjectivity of scoring resulted in a rather low Examination-reliability.

(5) A value of .65 was seen to be a reasonable estimate of the average reliabilities of the Examinations.

(6) One type of error only, that due to inconsistency of

different scorers, had its effect in reducing the reliabilities of the different tests by .02 to .40. The average reliability of the Examinations in general was reduced by .05 to .12.

C H A P T E R 3

EXPERIMENTING WITH OBJECTIVE TESTS

I. The Significance and Scope of the Experiment

In the preceding chapter it was shown that public examinations in Jordan have certain limitations and deficiencies the causes of which were seen to be ascribable mainly to the preparation and scoring of the conventional essay-type examinations in use.

Two remedies for these limitations and deficiencies are possible: improving the essay examinations themselves, and making use of standardized objective tests.

Recently, there have been attempts to improve the methods of Examinations by emphasizing "strict controls" over the setting of questions, the administration of Examinations, and the scoring of papers. It has already been shown that these controls could not eliminate errors of scoring nor could they make Examinations as satisfactory as they should be when rigorously judged for their validity, reliability, and practicability.

On the other hand, no standardized procedures of educational evaluation have yet been developed in Jordan.

To investigate the possibility of using standardized tests in public examinations, this writer constructed four objective tests on four topics of the Preparatory Examination, namely Mathematics,

Sciences, English Language, and Arabic Language. These tests were administered to a sample of students who had taken the official Preparatory Examination. They were then analyzed, using standard procedures, in a manner to be illustrated in detail in this chapter.

The Preparatory Examination level was chosen as a convenient level for investigating the possibility of using objective tests for the following reasons:

(1) The Preparatory Examination has **well-defined functions** that lend themselves most satisfactorily to objective testing. These functions are mainly, "testing a standard level of achievement" and "selection of students". The significance of these functions at this examination level is that the educational outcomes that need to be evaluated can be defined in terms of objectives that fit more satisfactorily in a well-designed objective test than in the kinds of essay examinations that are being used.

(2) Follow-up studies necessary for validation of the tests are facilitated by the fact that most students stay in the secondary cycle for three years after they have passed the Examination.

(3) The age level of students assures to a large extent efficient functioning of the tests. The students have reached a stage of readiness which makes it easy for the examiner to translate their behavior into objective items. Moreover, they can easily understand instructions for answering and can without much difficulty respond to the different types of items in a test.

(4) The function of tests designed for this level does not need to be the testing of achievement only. Since they can be used for selection and classification purposes, it would be useful to include items that test the aptitude for achievement. This allows for flexibility in the content and scope of a test, although it requires more careful planning. This also means that well-prepared objective tests can be used exclusively to serve the purposes of the Preparatory Examination, and that there would be no need to use in this Examination essay questions even on composition, because they would spoil its qualities as a measuring instrument more than improve its function for evaluation. If it is claimed that certain functions are better appraised by essay tests, it is better to split these functions into elements or components that lend themselves to item construction. The appraisal of these functions can be accomplished to a large degree and with great accuracy in terms of carefully identified elements or components.

It must be mentioned that the tests that have been constructed and tried out are basically achievement tests. The testing of aptitude would have required some extensive study of the different school programs so as to have a clear idea of the specific student behaviors that need be appraised. Therefore , the testing of aptitude was not considered when the tests were prepared. Moreover, the tests were meant to be equivalent forms (not parallel forms) to

the official examinations, so as to illustrate the contrast between the two methods of testing, essay examinations and objective tests. These tests can also be useful for future testing programs in supplying convincing evidence that the material represented in them could be successfully formulated into objective items, that the achievement in other subjects can also be tested objectively, that the consistency of scoring objective tests is very reliable, that validity and reliability are most likely to be high when the tests are carefully prepared, and that the technique of production, administration, and scoring of objective tests saves much administrative and clerical work and reduces the expenses to the minimum.

The processes involved in the whole testing experiment can be grouped under the following headings which will be discussed seperately in the following pages:

1. Planning the tests,
2. Preparing the tests,
3. Trying out the tests,

and 4. Evaluating the tests.

II. Planning the Tests

Test-planning consisted of making decisions on (a) the function of the test, (b) the course-objectives that can be included in objective items, and (c) the sources of material.

(a) The Function of the Tests:

Since the function of the Preparatory Examination in its present form is predominantly the testing of students' achievement in the prescribed course of study, it was decided to construct objective tests that have the same function so as to find out which of the two, the official essay tests or the objective ones, performs the same function most satisfactorily. Thus the two tests on any topic, the official test and the objective one, can be considered equivalent forms in that both of them test the achievement in the same course content, but they are not parallel forms because the questions are entirely different.

(b) The Course-Objectives:

Course-objectives were identified as follows:

(1) by reading the list of course-objectives specified in the syllabus and choosing from amongst them those objectives that can be adapted to objective items;

(2) by reading the prescribed text-book and classifying the subject-matter under certain objectives so that the part of the subject-matter and the corresponding objective can fit well into objective items;

(3) by using the general knowledge, common sense, and personal experience of the writer.

The weakness of this procedure is that subject-specialists and books other than the text and the syllabus were not consulted.

The main objectives that were included in the tests are listed below:

Mathematics:

1. Knowledge of mathematical concepts, relationships, theories, laws, and principles (verbal and mathematical).
2. Knowledge of mathematical conventions, technical terms, symbols dimensions, etc.
3. Ability to do abstract thinking, to understand the mathematical logic, and to use logical reasoning.
4. Ability to reason quantitatively and symbolically, to use symbolic relations, to perform numerical operations, and understand and use diagrams.
5. Ability to distinguish between false and correct statements.

Science:

1. Knowledge of facts, technical terms, symbols, units, dimensions, etc.
2. Knowledge of relationships, theories, laws, and principles (verbal and mathematical).
3. Ability to use the methods of science, to define a problem, to interpret data, to make a hypothesis and draw a conclusion,

to express in scientific terms, and to distinguish between facts and false statements.

4. Ability to reason quantitatively and symbolically, to use symbolic relations, to perform numerical operations, and to understand and use diagrams.
5. Ability to apply principles and relationships to novel situations.

English Language:

1. Knowledge of the basic concepts of the structure and grammar of the language.
2. Ability to understand written language and to interpret it.
3. Mastery of vocabulary and of proper use of words.
4. Ability to understand and appreciate the literary work and to interpret it.
5. Ability to distinguish between false statements and correct statements.
6. Ability to translate into one's own language, and visa versa.

Arabic Language:

Objectives 1 to 5 for English Language also apply for Arabic Language.

6. Knowledge of the outstanding features in the history of Arabic Literature, of famous literary works, and of the life history of famous literary characters.

(c) The Sources of Material:

Here, too, the choice of sources of material for the test items was determined by the idea of having the objective tests equivalent in content to the official tests. In the latter there was almost complete reliance on the syllabus and the text. The same sources were used for writing the objective tests; only for the English test some other books on grammar were consulted. However, only the comprehension paragraphs in both the English and Arabic tests, were taken directly from the texts, otherwise, the exact reproduction of statements or examples from the texts was scrupulously avoided.

III. Preparing the Tests

The preparing of each test proceeded as follows:

(1) The syllabus and the textbook were carefully studied. The subject -matter was classified into topics and sub-topics. A test "blue print" was prepared. This consisted of allocating items to topics and to objectives. The total number of items on each test was determined by two factors, the degree of complexity of material proposed to fill the items and the time proposed for answering the test. All the tests were assigned the same time; fifty minutes was considered a convenient time to start with, on the assumption that on first trial of the test, the time will be adjusted so that almost all students have enough time to attempt all the items.

(2) Preliminary draft of the test items was prepared. More items were included in the preliminary draft than would be needed in the final form. In constructing the test-items, the type of test-item that was favored was the multiple-choice type. Other types were used only when it was very difficult to cast the question in the multiple-choice form. And only in the Arabic Language test other types such as the True-False and the Completion types were used with the multiple-choice type. In the other three tests, the items were exclusively of the multiple-choice type.

(3) The items in the preliminary draft were revised, the wording was checked, and the best items were selected. The items selected were arranged in categories, with all items of the same type grouped together. The items in the same category were arranged in order of estimated difficulty having the easiest ones at the beginning. The correct answers in the lists of options were rearranged so that they looked randomly distributed and so that no particular order of correct answer-position has strong predominance over the others.

(4) General instructions for answering the test were situated at the openning page of the test booklet. Special instructions for guessing were included in the general instructions. Also there were instructions at the beginning of each set of items. The two tests on English and Arabic were organized so that the answers are checked, according to the instructions, on the test booklet itself. The other two tests on Mathematics and Science were provided with seperate answer sheets. Clear instructions for checking the answer sheet, with worked examples, were given in the general instructions.

(5) The tests were typed and duplicated in presentable form.[1] Copies of the same duplicates given to students are given in the Appendix.

---

[1]Before these tests were duplicated, they were read by the subject-inspectors in the Min. of Ed. This was requested by the Under Sec. of State upon asking him for permission to administer the tests in the Ministry's schools. The inspectors did not propose any changes in the tests.

IV. Trying out the Tests

The tests were administered to 111 students who have already passed the Preparatory Examination and who were coming for the first week to the first class in the secondary cycle. No special method has been used in selecting these students except that they were easily within reach. They belonged to two secondary schools in Amman, Hussein College and Hussein Secondary School. In the first school, 50 students were taken out of about 85 in the same class. In the second school, 61 students were taken out of about 90 in the same class. In either school the selection of students consisted of simply collecting the required number without any regard to their classification or their educational background. The tests were administered in each school seperately, but precautions were taken so that between the administrations of the same tests in the two schools the lapse of time was not sufficient for students in the two schools to get in touch with each other.

The school administration cooperated in arranging and supervising the students during the administration of the tests.

To insure earnest participation of the students, they were urged to answer the tests as seriously as in an examination. They were told that the tests would show how much each one of them remembered of his preparatory education and that valuable books will be offered as prizes to the one who ranks the first in any of

the tests. It was noticed that the students were interested to answer "the puzzle-like" tests, and that there was a sense of competition among them. Later, the papers were scored, scores were changed into percentile grades and handed to the headmasters with the prizes for the top four students.

In scoring the papers, final scores were computed by means of the following formula which corrects for chance success, thus minimizing the differences between students who guess much or little:

$$S = R - \frac{W}{n - 1}$$

where:

S = score corrected for guessing.
R = number of items answered correctly.
W = number of items answered incorrectly.
n = number of answer choices for an item.

V. Evaluating the Tests

(1) Validity:

No validation with a criterion has been made. The scholastic achievement, of the students who were tested, in the secondary cycle; or their achievement in the General Secondary Examination would serve as a criterion in case the tests are to be validated.

On the other hand, the tests can be considered to have content validity to the extent that (a) the planning of the tests and (b) the use of the prescribed material for the course, as described earlier in this chapter, are judged satisfactory for valid representation of the course content.

(2) Reliability:

An estimate of "single test" reliability is computed by means of the Kuder-Richardson Formula 21:[1]

$$r_{tt} = \frac{n}{n-1}\left[1 - \frac{M_t\left(1 - \frac{M_t}{n}\right)}{S_t^2}\right]$$

where:

r_{tt} = estimate of single test reliability,
n = number of items in the test,
M_t = mean score of the group,
S_t = standard deviation of the test.

---

[1] R. L. Thorndike and E. Hagen, Measurement and Evaluation in Psychology and Education, p. 181, New York, London; John Wiley & Sons, Inc., 2nd edition 1961.

The estimate of reliability computed for each of the four tests is given under "The Results of Test Analysis" in the following pages.

(3) Item Analysis:

Two indexes are obtained concerning each item, (a) the difficulty of the item and (b) the discrimination index of the item. The computation of item difficulty and discrimination is based on the responses made by individuals in the upper and lower 27% of the groups.[1]

(a) Item difficulty is the proportion of individuals in the combined upper and lower 27% groups who answer the item correctly.

(b) Item discrimination refers to how well an item discriminates between students who rank high and students who rank low on the test as a whole.[2] The method that will be followed in demonstrating item discrimination is the correlation method. The general principle in this method is to compute a correlation coefficient that shows the relationship of the responses to an item to the total score.[3] This means calculating a correlation coefficient for every item, which is a long and tedious process. This process has been simplified by J. C. Flanagan by computing a chart in which the correlation coefficients are read directly when one enters the chart

---

[1]The choice of 27% groups is after T. L. Kelly; adapted from N. M. Downie and R. W. Heath, Basic Statistical Methods, p. 203, Harper &Brothers, Publishers, New York, 1959.

[2]C. C. Ross, Measurement in Today's Schools, p. 124, New York, Prentice-Hall, Inc., 2nd edition, 1953.

with the percent answering the item correctly in the upper and lower 27% of the papers. The values of discrimination indexes of items in the present tests are based on Flanagan's chart;[4] a point estimate is approximated by using a table adapted from Flanagan's Original Table for the course in Education 228, "Theory and Methods of Testing" at the American University of Beirut.

The Results of Test Analysis

In the following pages, the results of test-analysis are presented as follows:

(1) Frequency distribution of scores on each test, with mean and standard deviation.

(2) The coefficient of reliability (r_{tt}) computed by means of the Kuder-Richardson Formula 21.

(3) A table showing the percentages of individuals in the top and bottom 27% groups answering each option in each item. (It should be noted that 100% of the top or bottom group is equal to 30 individuals, or 27% of 111.)

(4) A table showing percent difficulty and discrimination index for every item.

(5) A scatter diagram in which item difficulty (D) is plotted against discrimination index (r). This diagram pictures all the test items in such a way that it becomes easy to see what items can be considered to be functioning satisfactorily and what items

---

[3] N. M. Downie and R. W. Heath, Op. cit, p. 203.

[4] Ibid, p. 174

cannot.

Since (r) is supposedly based on a sample of 60 (the top and bottom groups), a test of significance will show that an r of .25 or more is significant at the .05 level. Hence a straight line drawn on the scatter diagram at $x \geqslant .25$ will isolate to its left all items that can be considered to be of unsatisfactory discrimination indices. An item near this line, lower than .25 in reliability, has, nevertheless, a good probability of being moderately operative in furthering the measurement produced by the test as a whole. Many test makers use some of these items in situations where their difficulty indices are particularly useful for bringing the total difficulty toward a desired level, or where too few items are found to the right of the line.

On the other hand, the scatter diagram will show on its upper end the items which can be considered too easy and on its lower end the items which can be considered too difficult. The item number is indicated at the plotted point corresponding to it so as to make it easy to refer to the item in the test for further consideration.

As the critical study of each item is beyond the scope of this thesis, it suffices to mention some of the considerations according to which a decision can be made:

(a) Evidently, an item with high discrimination index is kept. Items with very low discrimination index are to be reconsidered: modified, reworded, discarded, etc.

(b) In most achievement tests an average item difficulty of 50% is most desirable. Also it is desirable that a considerable number of items cluster about the 50% difficulty level, with some spreading as far as .10 and .90. There are special situations where a high average percent difficulty, and other situations where a low percent difficulty, is desired. This depends upon the function which the test is designed to perform.

(c) In considering a particular item, a study of the responses to all of its options must be made. The views of examinees, if they can be unreservedly secured, about the different items can be enlightening about the weaknesses of the items and how they should be modified.

(d) It is very useful to consult with subject-specialists and with persons experienced in item construction.

Mathematics Test

Number of test items: 30

Number of participants: 111 (50 first secondary students at Hussein College - Amman, and 61 first secondary students at Hussein Secondary School - Amman.)

Maximum time allowed: 60 minutes

Date of administration: September 12,1962.

Scores obtained: 28, 24, 22, 21(2), 20(2), 19(3), 18(5), 17(2), 16(2), 15(3), 14(4), 13(7), 12(7), 11(11), 10(12), 9(13), 8(4), 7(6), 6(5), 5(2), 4(3), 3, 2(5), 1(3), 0(6)

(Numbers in parentheses show number of students obtaining that score)

Mean = M_t = 10.34

Standard deviation = S_t = 5.69

Reliability = r_{tt} = .818

Mathematics Test

Item Analysis

(Numbers are percentage of Top and Bottom groups; $N_{top} = 30$, $N_{bottom} = 30$; 30 = 27% of 111. Underlining identifies correct response; 0 = omitted.)

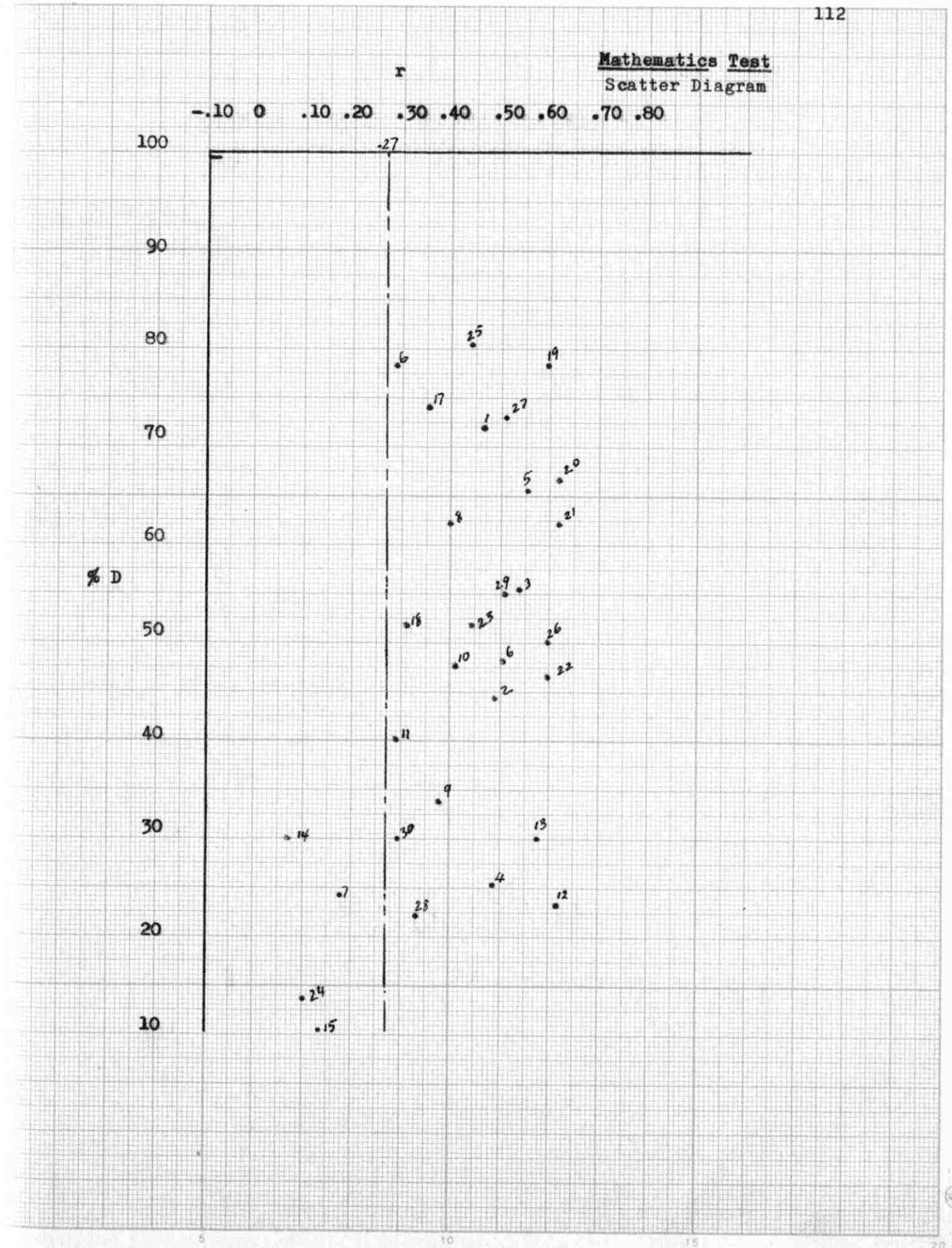
	A	B	C	D	0
1.	0	0	10	90	0
	4	20	20	53	3
2.	67	0	7	20	6
	20	3	34	23	20
3.	7	3	83	0	7
	47	10	30	3	10
4.	10	24	43	13	10
	47	20	7	16	10
5.	7	3	0	90	0
	43	10	4	40	3
6.	10	3	73	4	10
	10	43	23	7	17
7.	13	47	0	30	10
	0	47	10	17	26
8.	3	80	14	0	3
	3	43	17	27	10
9.	3	0	17	50	30
	30	23	10	17	20
10.	67	13	7	0	13
	27	30	33	3	7
11.	53	7	3	37	0
	27	17	17	33	6
12.	0	7	20	43	30
	13	33	27	3	24
13.	53	10	13	7	17
	7	17	7	43	26
14.	20	33	30	0	17
	10	27	23	17	23
15.	13	13	17	20	37
	40	7	23	3	27
16.	0	87	0	10	3
	3	67	3	20	7
17.	3	3	87	3	4
	13	10	60	10	7
18.	67	27	3	3	0
	37	53	7	0	3
19.	97	3	0	0	0
	60	30	3	0	7
20.	3	0	93	0	4
	17	17	40	20	6
21.	90	3	3	0	4
	33	13	44	0	10
22.	7	7	3	80	3
	40	7	37	13	3
23.	3	73	0	17	7
	13	30	13	30	14
24.	13	17	13	20	37
	3	14	13	50	20
25.	3	3	94	0	0
	20	7	67	3	3
26.	7	10	3	80	0
	36	7	27	20	10
27.	0	7	0	93	0
	10	27	7	53	3
28.	13	33	27	24	3
	17	10	23	37	13
29.	0	13	80	7	0
	20	23	30	20	7
30.	43	43	0	10	4
	33	17	20	23	7

Mathematics Test

Item Analysis

(D = percent difficulty; r = discrimination index)

	% D	r		% D	r
1.	72	.46	16.	77	.28
2.	44	.48	17.	74	.35
3.	56	.54	18.	52	.31
4.	25	.49	19.	78	.59
5.	65	.56	20.	66	.61
6.	48	.50	21.	62	.61
7.	24	.18	22.	46	.66
8.	62	.40	23.	52	.43
9.	34	.38	24.	13	.00
10.	47	.41	25.	80	.44
11.	40	.28	26.	50	.60
12.	23	.61	27.	73	.51
13.	30	.57	28.	22	.33
14.	30	.07	29.	55	.51
15.	10	.14	30.	30	.30


Mathematics Test
Scatter Diagram
r
-.10 0 .10 .20 .30 .40 .50 .60 .70 .80
.27
100
90
80
70
60
50
40
30
20
10
% D
25
6
19
17
1
27
5
20
8
21
29
3
18
23
26
10
6
22
2
11
9
14
30
13
4
7
28
12
24
15
5
10
15
20

Science Test

Number of test items: 40

Number of participants: 111 (50 first secondary students at Hussein College, Amman; and 61 first secondary students at Hussein Secondary School - Amman.)

Maximum time allowed: 50 minutes

Date of administration: September 12, 1962.

Scores obtained: 31(2), 28, 27, 26, 24(2), 23(6), 22(5), 21(4), 20(13), 19(5), 18(6), 17(8), 16(7), 15(13), 14(7), 13(3), 12(7), 11(5), 10, 9(4), 8(4), 7(3), 5(2).

(Numbers in parentheses show number of students obtaining that score.)

Mean = M_t = 16.55

Standard deviation = S_t = 5.37

Reliability = r_{tt} = .68

Science Test

Item Analysis

(Numbers are percentages of Top and Bottom groups; $N_{top} = 30$, $N_{bottom} = 30$; 30 = 27% of 111. Underlining identifies correct response; 0 = omitted.)

	A	B	C	D	O
1.	0	10	0	90	0
	13	24	3	60	0
2.	0	0	100	0	0
	0	7	93	0	0
3.	3	73	17	0	7
	7	7	73	6	7
4.	4	3	93	0	0
	17	3	53	20	7
5.	93	4	0	0	3
	60	13	7	17	3
6.	3	97	0	0	0
	30	50	13	3	4
7.	20	7	50	10	13
	20	23	40	13	4
8.	20	30	27	16	7
	7	27	56	3	7
9.	0	0	97	3	0
	17	16	60	0	7
10.	83	0	10	0	7
	37	10	23	26	3
11.	7	0	93	0	0
	10	13	63	7	7
12.	3	37	10	40	10
	10	47	10	26	7
13.	0	0	100	0	0
	7	10	80	3	0
14.	17	0	0	83	0
	53	7	0	37	3
15.	3	0	0	97	0
	0	7	0	93	0

	A	B	C	D	O
16.	93	3	4	0	0
	90	10	0	0	0
17.	74	23	0	0	3
	83	7	7	0	3
18.	3	93	0	4	0
	13	60	3	24	0
19.	87	10	0	-	3
	57	13	27	-	3
20.	80	0	17	0	3
	77	7	10	3	3
21.	3	67	23	3	4
	23	27	47	3	0
22.	3	70	7	0	3
	7	27	50	16	0
23.	33	0	40	23	4
	30	10	37	23	0
24.	3	97	0	-	0
	13	73	10	-	4
25.	77	16	7	-	0
	40	50	7	-	3
26.	0	20	77	0	3
	10	43	37	3	0
27.	10	3	73	14	3
	0	3	33	53	10
28.	7	0	0	93	0
	7	7	3	77	6
29.	0	17	23	53	6
	13	3	17	53	14
30.	40	3	33	20	4
	30	30	30	3	7

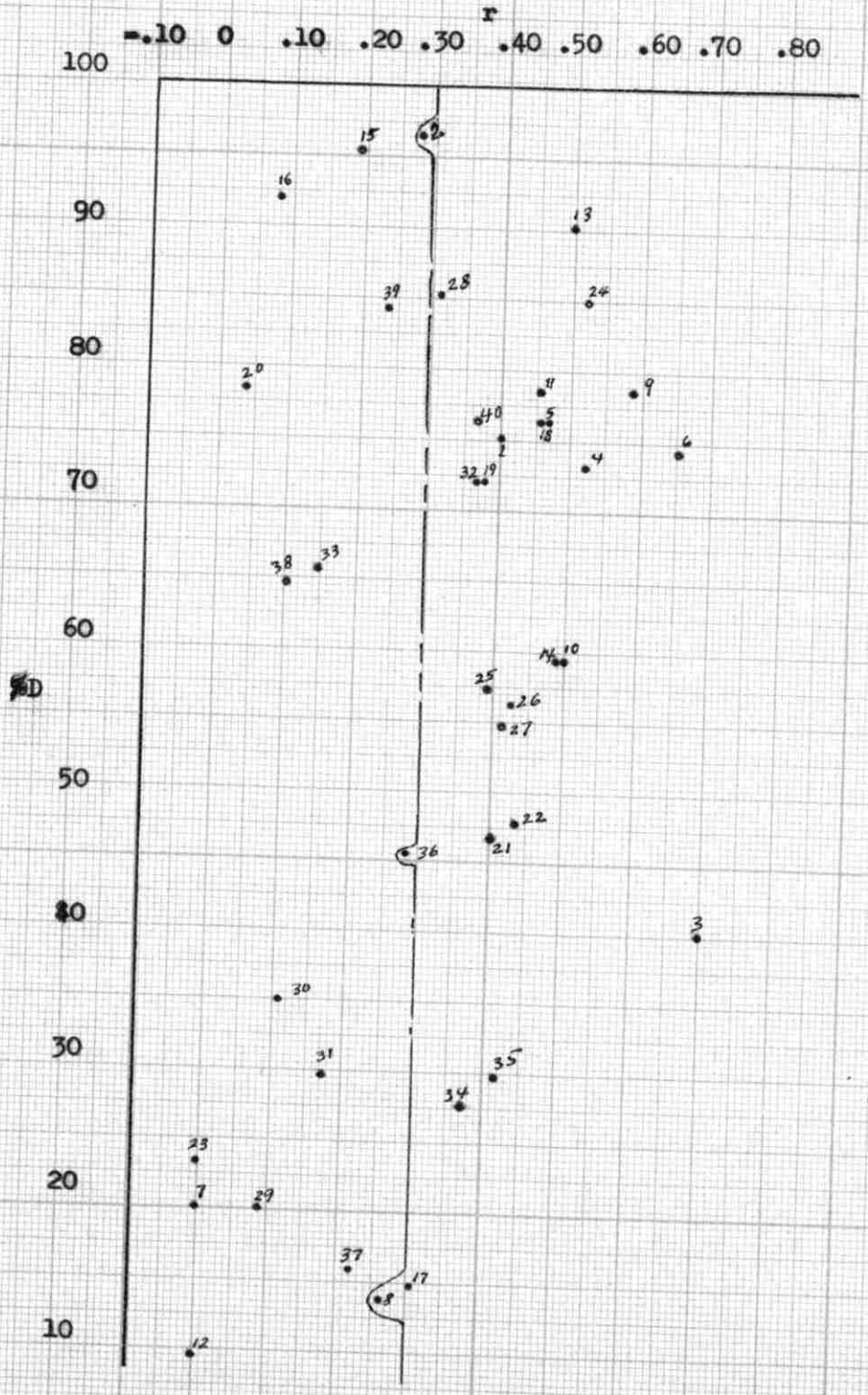
E
0
7

	A	B	C	D	O
31.	40	37	23	0	0
	33	23	34	7	3
32.	3	87	0	3	7
	20	57	0	13	10
33.	73	3	7	10	7
	60	13	17	0	10
34.	24	10	43	13	10
	43	17	13	13	14
35.	27	16	47	10	0
	27	10	13	40	10
36.	17	3	10	60	10
	20	3	24	33	20
37.	10	17	20	23	30
	27	40	13	10	10
38.	10	3	70	14	3
	13	3	60	17	7
39.	90	3	0	7	0
	77	7	13	0	3
40.	0	90	7	0	3
	10	63	13	10	4

(D = percent difficulty; r = discrimination index)

	% D	r		% D	r
1.	75	.40	21.	47	.40
2.	96	.27	22.	48	.43
3.	40	.69	23.	23	.00
4.	73	.52	24.	85	.51
5.	76	.44	25.	58	.39
6.	74	.65	26.	57	.42
7.	20	.00	27.	55	.41
8.	14	.26	28.	85	.31
9.	78	.59	29.	20	.09
10.	60	.49	30.	35	.11
11.	78	.46	31.	30	.17
12.	10	.00	32.	72	.37
13.	90	.49	33.	66	.15
14.	60	.49	34.	28	.37
15.	95	.19	35.	30	.41
16.	92	.08	36.	46	.28
17.	15	.30	37.	16	.22
18.	76	.44	38.	65	.11
19	72	.38	39.	84	.22
20.	78	.04	40.	76	.37

Science Test
Scatter Diagram



English Language Test

Number of test items: 50

Number of participants: 111 (50 first secondary students at Hussein College -Amman, and 61 first secondary students at Hussein Secondary School -Amman.)

Maximum time allowed: 50 minutes.

Date of administration: September 11, 1962.

Scores obtained: 49, 45, 39(3), 36(5), 35, 34(3), 32(3), 31, 30(6), 29(2), 28(2), 27(8), 26(6), 25(2), 24(5), 23(3), 22(6), 21(2), 20(3), 19(4), 18, 17(5), 16(9), 15(3), 14(2), 13(3), 11(2), 9(4), 8, 6(5), 5(3), 4(2), 3, 2(2).

(numbers in parentheses show number of students obtaining that score.)

Mean = M_t = 21.68

Standard deviation = S_t = 10.30

Reliability = r_{tt} = .904

English Language Test

Item Analysis

(Numbers are percentages of Top and Bottom groups; $N_{top} = 30$, $N_{bottom} = 30$; 30 = 27% of 111. Underlining identifies correct response; 0 = omitted.)

	A	B	C	D	O
1.	0	0	100		0
	10	0	<u>90</u>		0
2.	10	90	0		0
	17	<u>80</u>	0		3
3.	97	0	3		0
	<u>67</u>	10	20		3
4.	100	0	0		0
	<u>93</u>	3	3		0
5.	0	0	100		0
	27	13	<u>60</u>		0
6.	27	63	10		0
	60	<u>30</u>	3		7
7.	7	3	83		7
	27	33	<u>30</u>		10
8.	63	17	20		0
	<u>17</u>	30	40		13
9.	3	87	10		0
	17	<u>23</u>	53		7
10.	97	3	0		0
	<u>67</u>	23	10		0
11.	33	43	20		3
	<u>30</u>	37	30		3
12.	0	77	20		3
	13	<u>47</u>	37		3
13.	67	0	20	10	3
	<u>27</u>	13	20	33	7
14.	26	57	7	10	0
	24	<u>13</u>	53	10	0
15.	3	4	93	0	0
	33	30	<u>37</u>	0	0
16.	100	0	0	0	0
	<u>83</u>	14	3	0	0
17.	73	14	3	10	0
	<u>43</u>	17	23	14	3
18.	10	3	70	14	3
	17	33	<u>17</u>	30	3
19.	3	10	4	83	0
	30	23	7	<u>40</u>	0
20.	3	10	60	27	0
	13	7	<u>40</u>	40	0
21.	14	0	3	83	0
	40	26	7	<u>20</u>	7
22.	40	3	47	7	3
	33	27	<u>10</u>	23	7
23.	10	0	0	87	3
	33	10	7	<u>50</u>	0
24.	57	33	0	3	7
	<u>50</u>	20	14	13	3

English Language Test

Item Analysis

	A	B	C	D	O
25.	83	4	3	7	3
	10	60	10	7	13
26.	3	0	30	67	0
	34	10	53	0	3
27.	7	77	6	7	3
	14	33	50	0	3
28.	17	10	63	7	3
	23	57	13	7	0
29.	7	3	77	3	10
	27	13	37	16	7
30.	43	20	14	20	3
	10	36	27	17	10
31.	20	30	17	20	13
	23	27	17	17	16
32.	0	0	100	0	0
	13	7	67	10	13
33.	93	0	4	3	0
	50	13	10	20	7
34.	3	0	4	93	0
	13	7	16	47	10
35.	0	60	3	37	0
	7	20	20	46	7
36.	3	87	7	0	3
	40	30	7	3	20
37.	57	20	10	7	6
	23	10	30	13	24
38.	0	0	97	3	0
	33	10	37	10	10
39.	10	3	87	0	0
	13	10	47	20	10
40.	7	80	0	10	3
	20	27	13	23	17
41.	20	7	0	73	0
	13	23	10	40	13
42.	0	100	0	0	0
	23	40	23	13	0
43.	0	0	7	93	0
	17	7	27	50	0
44.	100	0	0	0	0
	77	20	0	0	3
45.	100	0	0	0	0
	70	13	7	10	0
46.	7	76	0	17	0
	22	27	17	27	7
47.	3	26	47	17	7
	20	20	17	30	13
48.	77	3	10	3	7
	17	10	30	16	27
49.	37	0	7	43	13
	17	17	16	30	20
50.	30	3	57	3	7
	30	3	33	17	17

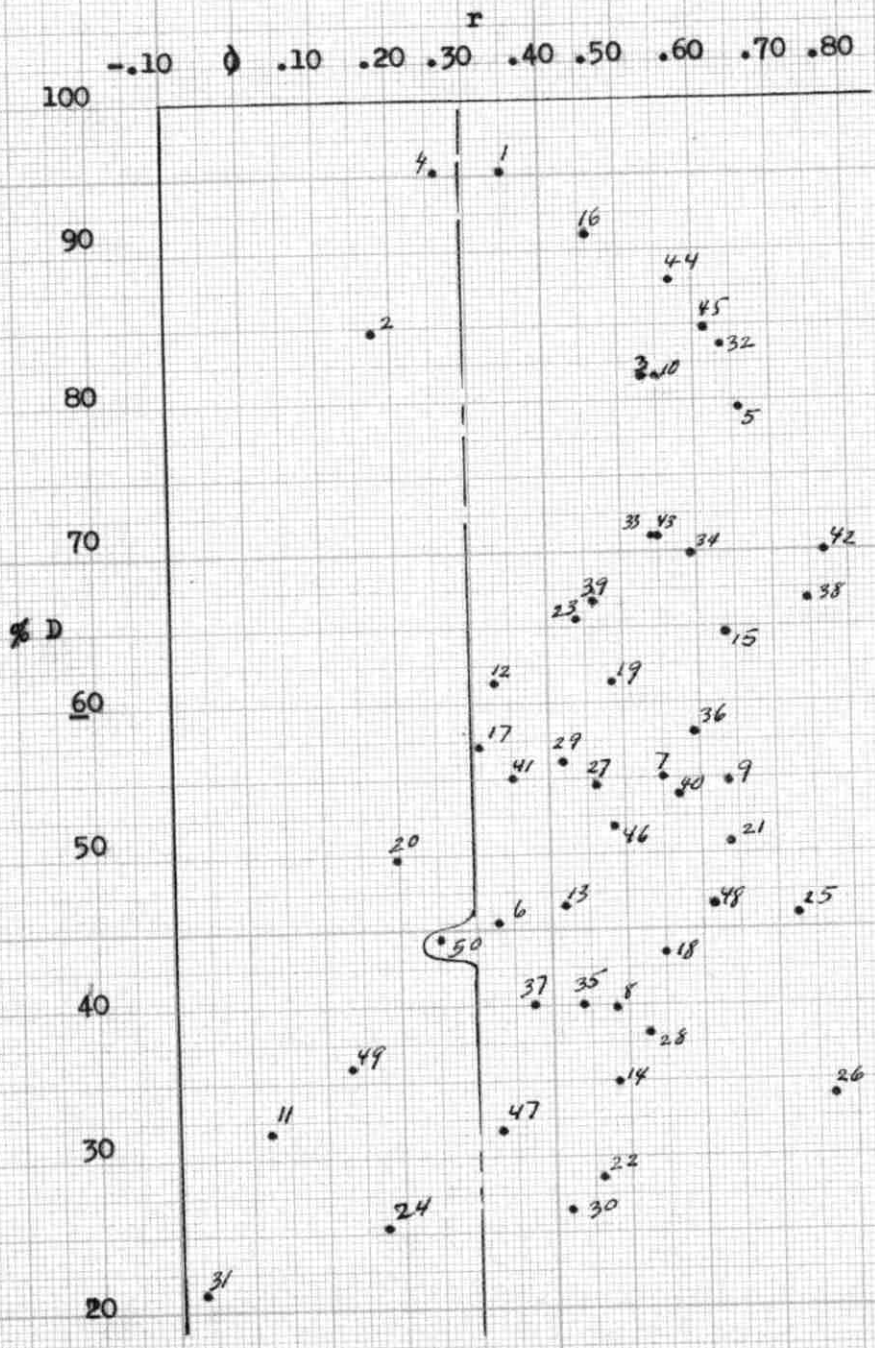
English Language Test

Item Analysis

	D	r		D	r
1.	95	.35	26.	34	.76
2.	85	.18	27.	55	.45
3.	82	.54	28.	38	.53
4.	96	.25	29.	57	.42
5.	80	.66	30.	26	.42
6.	46	.34	31.	22	- .04
7.	56	.54	32.	84	.64
8.	40	.48	33.	72	.54
9.	55	.63	34.	70	.59
10.	82	.55	35.	40	.43
11.	32	.03	36.	58	.58
12.	62	.32	37.	40	.36
13.	47	.41	38.	67	.73
14.	35	.48	39.	67	.45
15.	65	.64	40.	54	.57
16.	91	.46	41.	56	.34
17.	58	.31	42.	70	.77
18.	44	.54	43.	72	.54
19.	62	.48	44.	88	.56
20.	50	.20	45.	85	.62
21.	51	.64	46.	52	.49
22.	28	.46	47.	32	.34
23.	66	.44	48.	47	.60
24.	26	.17	49.	36	.14
25.	46	.72	50.	45	.25

English Language Test

Scatter Diagram



Arabic Language Test

Number of test items: 58

Number of participants: 111 (50 first secondary students at Hussein College -Amman, and 61 first secondary students at Hussein Secondary School - Amman.)

Maximum time allowed: 50 minutes

Date of administration: September 11, 1962

Scores obtained: 52, 50(2), 49, 47, 46(5), 45, 44(5), 43(3), 42, 41(3), 40(4), 39(5), 38(5), 37(9), 36(9), 35(8), 34(9), 33(9), 32(4), 31(4), 30(4), 29(6), 28(2), 27(2), 26(3), 25(2), 23,22, 21.

(Numbers in parentheses show number of students obtaining that score.)

Mean = M_t = 35.78

Standard deviation = S_t = 6.38

Reliability = r_{tt} = .675

Arabic Language Test

Item Analysis

(Numbers are percentages of Top and Bottom groups; $N_{top} = 30$, $N_{bottom} = 30$; 30 = 27% of 111. Underlining identifies correct response; 0 = omitted.)

	A	B	C	D	O
1.	0	3	93	0	4
	7	7	70	0	16
2.	80	7	10	0	3
	40	23	23	7	7
3.	3	90	0	3	4
	17	53	3	10	17
4.	43	33	4	3	17
	20	57	0	10	13
5.	0	0	0	100	0
	0	3	0	90	7
6.	0	97	3	0	0
	0	80	10	3	7
7.	83	7	3	0	7
	60	13	3	4	20
8.	23	14	23	30	10
	17	10	30	20	23
9.	30	0	10	53	7
	20	3	30	37	10
10.	13	3	77	0	7
	33	37	27	0	3
11.	3	0	83	7	7
	30	7	27	23	13
12.	3	97	0	0	0
	17	40	20	3	20
13.	0	0	87	13	0
	7	3	33	30	27
14.	100	0	0	0	0
	80	0	3	7	10
15.	0	97	0	0	3
	20	37	27	3	13

	A	B	C	D	O
18.	0	0	100	0	0
	0	0	90	7	3
19.	0	100	0	0	0
	0	50	40	7	3
20.	97	0	3	0	0
	77	7	7	6	3
21.	0	7	93	0	0
	7	3	87	3	0

	T	F	O
22.	0	100	0
	7	93	0
23.	100	0	0
	90	10	0
24.	0	100	0
	3	97	0
25.	0	100	0
	13	87	0
26.	97	3	0
	83	17	0
27.	3	97	0
	17	83	0
28.	83	17	0
	73	23	4
29.	0	97	3
	7	93	0
30.	0	100	0
	3	97	0
31.	100	0	0
	87	13	0

Arabic Language Test

Item Analysis

	A	B	C	D	O
16.	13	0	0	87	0
	7	3	7	<u>76</u>	7
17.	0	0	0	100	0
	0	3	0	<u>97</u>	0

	T	F	O
32.	100	0	0
	<u>97</u>	3	0

	R	W	O
33.	90	7	3
	<u>83</u>	17	0
34.	57	27	16
	<u>50</u>	47	3
35.	83	10	7
	<u>57</u>	33	10
36.	93	7	0
	<u>73</u>	23	4
37.	60	30	10
	<u>27</u>	70	3
38.	47	47	6
	<u>27</u>	67	6
39.	43	54	3
	<u>27</u>	66	7
40.	40	53	7
	<u>23</u>	77	0

	A	B	O
41.	80	17	3
	<u>37</u>	60	3
42.	0	100	0
	0	<u>100</u>	0
44.	0	100	0
	13	<u>87</u>	0
45.	97	3	0
	<u>87</u>	13	0
46.	40	60	0
	46	<u>47</u>	7
47.	3	97	0
	0	<u>97</u>	3
48.	87	13	0
	<u>77</u>	20	3
49.	97	3	0
	<u>73</u>	23	4
50.	13	83	4
	27	<u>73</u>	0

	R	W	O
51.	83	17	0
	<u>77</u>	20	3
52.	83	14	3
	<u>53</u>	23	24
53.	80	17	3
	<u>57</u>	27	16
54.	63	20	17
	<u>53</u>	10	37
56.	70	30	0
	<u>30</u>	70	0
57.	83	17	0
	<u>50</u>	50	0
58.	93	7	0
	<u>70</u>	30	0
59.	80	17	3
	<u>37</u>	63	0
60.	90	10	0
	<u>63</u>	33	4

Arabic Language Test

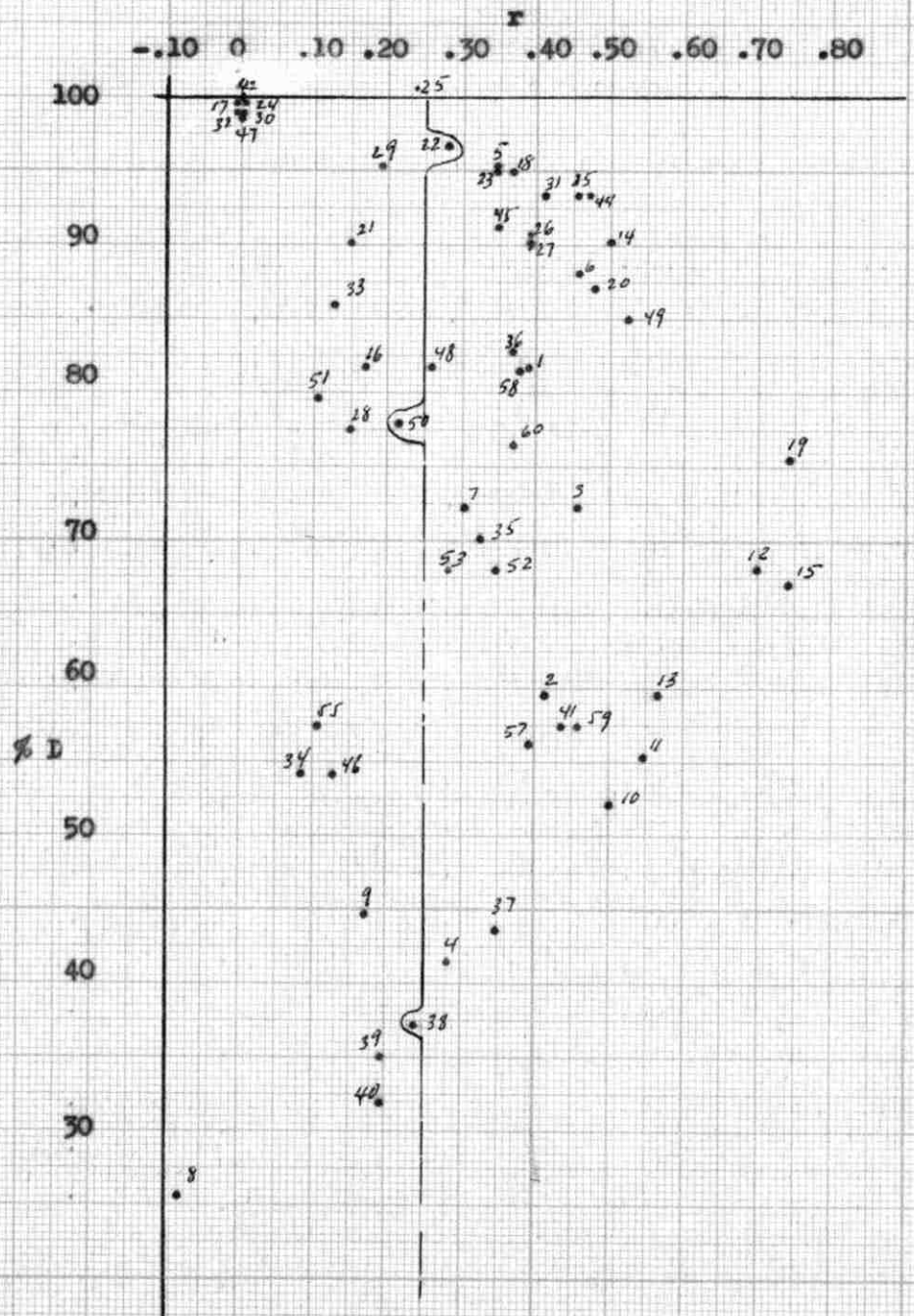
Item Analysis

(D = percent difficulty; r = discrimination index)

	D	r		D	r		D	r
1.	82	.38	21.	90	.15	41.	58	.44
2.	60	.42	22.	96	.27	42.	100	.00
3.	72	.46	23.	95	.35	44.	94	. 44
4.	42	.27	24.	98	.00	45.	92	.34
5.	95	.35	25.	94	.44	46.	54	.13
6.	88	.45	26.	90	.38	47.	97	.00
7.	72	.30	27.	90	.38	48.	82	.26
8.	26	- .09	28.	78	.14	49.	85	.51
9.	45	.16	29.	95	.19	50.	78	.22
10.	52	.50	30.	98	.00	51.	80	.09
11.	55	.56	31.	94	.39	52.	68	.35
12.	68	.70	32.	98	.00	53.	68	.27
13.	60	.55	33.	86	.13	55.	58	.10
14.	90	.49	34.	54	.07	56.	50	.40
15.	67	.73	35.	70	.31	57.	56	.38
16.	82	.16	36.	83	.36	58.	82	.38
17.	98	.00	37.	44	.34	59.	58	.45
18.	95	.35	38.	37	.22	60.	76	.37
19.	75	.74	39.	35	.18			
20.	87	.48	40.	32	.20			

Arabic Language Test

Scatter Diagram



C H A P T E R 4

COMPARISON OF RESULTS OF STUDY ON OBJECTIVE TESTS WITH RESULTS OF STUDY ON PUBLIC EXAMINATIONS

In the preceding chapters, the methods of public examinations were investigated and, within the scope of the investigation, certain defects have been identified. Objective tests were then tried, not only as a proposed remedy for these defects, but also in an attempt to provide some convincing evidence that the use of objective tests constitutes an improvement and progress in the methods of evaluation employed in public examinations.

To find out whether or not there is sufficient justification for the use of objective tests in public examinations and whether or not the use of objective tests in these examinations would signify any improvement or progress in method, the results of studies on public examinations and on objective tests are compared as illustrated below. The comparison is divided into three parts corresponding to the three criteria by which a measuring instrument is judged satisfactory, namely validity, reliability, and practicability.

I. Compatibility with Valid Measures:

That public examinations in Jordan are not very compatible with valid measures is confirmed by the following results from the investigation on General Examinations:

(a) The purposes of the Examinations were not well recognized by those who design the Examinations. The purposes in terms of course-objectives and specific student behavior were not very clear in the minds of those who set the questions.

In setting the questions, many of the objectives essential for teaching the course were ignored. While in the greatest part of the questions, the student is required to reproduce the material of the text, in some other parts he is required only to solve mathematical problems. Thus in a whole set of questions on one of the Examination topics, only one or two objectives are included. Testing the capacity to memorize is the predominant one.

(b) The function set for the Examinations is only partially achieved because of defects in test-planning, because the questions included relatively narrow sampling of course-content and student's knowledge, and because class instruction was to a large extent ignored.

That the objective tests that were tried out are likely to be more compatible with valid measures is supplemented by the "merits of the plan" followed in constructing the tests. Since

there was not the opportunity to test for validity, it is reasonable to assume for all practical purposes, that a plan which has been elaborately worked out by many authorities in the field of testing will most likely produce satisfactory results. This writer, although he is not a specialist on all the topics of the tests, tried, to his best knowledge and within the available time and resources, to carry out such a plan in the way that has been demonstrated in the preceding chapter. Objective tests can be as bad as essay tests if they are not carefully prepared. But there are recommended procedures that tend to make objective tests very valid measures. Hence if any shortcomings were identified in the tests prepared for the purpose of this thesis they do not condemn objective testing as an unvalid procedure, for better tests can be prepared by eliminating those shortcomings, by devoting more time and thinking to perfecting them, and by having subject-specialists contribute their valuable knowledge and experience.

In spite of all this, the least that can be said about these tests is that there have been some features in the plan followed that contribute favorably to their validity. Also it can be noticed that these features stand out as remedies for some of the defects in the General Examinations. Thus:

(a) The purposes of public examinations were identified in the whole system of education and in the official statement of educational policy. One of these purposes, "testing the standard

level of achievement" which the objective tests were designed to accomplish, could be translated into well-defined course-objectives and course-content.

(b) The items were constructed so that they would include a variety of objectives that are essential to the course of study. Testing the student's ability to reproduce the material of the text was considered of least importance. Admittedly, the textbook, and the syllabus, were the sources of material for the items. But while in most of the questions on the General Examinations, the student is required to reproduce an answer the way it is stated in the text, in the objective tests the items are structured in different ways so that different mental processes, and not only memory, are called upon to function.

(c) There was a much larger controlled sampling of course-content in the objective tests than in the General Examination-questions.

II. Compatibility with Reliable Measures:

The investigation on General Examinations has shown that these examinations suffer from the subjectivity of the questions and the subjectivity of scoring. It was shown that the subjectivity of scoring has caused two types of errors referred to as human errors. It was also demonstrated that these errors reduced grievely the reliabilities of the Examinations. The results of the investigation supplemented by findings from previous researches on

essay tests indicate that the average coefficient of Examination reliabilities is most likely less than .65.

On the other hand, the scoring of objective tests does not involve human errors. By virtue of the item structure, objective test-items do not include subjective or indefinite questions except when the rules recommended for item construction are not carefully applied.

In the experiment with objective tests, a reliability coefficient as large as .90 was obtained on one of the tests. On the other tests, the smallest coefficient obtained was .68. Within the limitations of the experiment these results are very encouraging. They indicate that better results can be obtained when conditions better than those under which the experiment was conducted are available. "Moreover, these reliabilities may be expected to increase substantially upon the elimination and revision of items with poor characteristics, and the addition of items similar to the most effective ones. It is one of the strong points of objective testing that the process of item analysis clearly identifies the weak items in a test and provides unmistakable signposts for the strengthening of its powers."[1]

---

[1] Professor F. Korf, Advisor of this thesis, Head of Office of Tests and Measurements, A.U.B., 1962-1963.

III. Compatibility with Practical Measures:

The validity and reliability of a test are usually considered of primary importance because they have to do with the accuracy of the test. However, there are other practical considerations which must be taken into account. The degree to which a test can be successfully applied without much expenditure of time and effort is a desirable quality referred to as practicability.

When objective tests were prepared, all those features that make a test of practical value were taken into account: well-organized lay-out of the items, clear instructions for answering, clear typing and neat duplication, minimum cost of paper and of reproduction. Some other features that make tests of practical value when standardized are that they can be used several times and they provide useful information for interpreting the results.

The practicability of General Examinations was not critically investigated, but it goes without saying that the impracticability associated with common essay examinations applies to General Examinations in Jordan as well. In general, essay examinations cannot be as practical as objective tests. Truly, the questions can be prepared quickly, but the same questions are scarcely used twice. The essay questions do not usually need specific instructions for answering as in an objective test, but many a time the student knows "what" to answer but does not know "how".

In the matter of time required for administration and scoring, a good objective test usually requires less than one minute per score point in administration, and less than one minute per twenty score points in scoring. The average essay test in Jordan requires about 135 minutes to administer, yields 100 score points, and takes about 30 minutes to score. An objective test yielding the same number of score points would do the same job, in a total of less than 80 minutes, or less than half the time , simultaneously eliminating human error in the scoring.

The expenditure in time and energy involved in scoring essay examinations can be illustrated by an example from the General Preparatory Examination. In the Balka District, the scoring of about 3,000 papers of a 2-hour test on English was accomplished by 25 scorers in about 10 days, working 6 hours per day. This means that, on the average, more than half an hour was needed for the scoring of one paper. If each scorer is paid J. Fils 35 for scoring one paper of 1-hour test,[1] then over J.Dinar 200 are paid for scoring all the papers of the English test. The preparatory Examination includes nine other 2-hour tests and is administered in six other districts. This means that the expenses of scoring the whole examination are equal to about 50 to 60 times the expenses of scoring the English test. A rough estimation would be about J.Dinars 10,000. Had the test on English been formulated into

---

[1] Jordan, Ministry of Ed., The System of the Gen. Prep. Cert. Exam. for 1959, (in Arabic).

objective items (a reasonable number of items could not be more than 100 in this case), a paper would not need more than five minutes to be scored when an answer sheet is used. On the assumption that the paper of the essay form needed thirty minutes to be scored, then both the time and cost of scoring the English test will be reduced by using the objective form to less than one sixth. The same reasoning applies to the other tests on the Examination with the result that the cost and time of scoring are very much reduced.

The actual justification of using objective tests is not because they save money, time, and energy only, but, as it has been illustrated by the whole comparison, because they can very well satisfy all the provisions that General Examinations are expected to fulfil.

The following comparative outline shows some of the basic advantages which objective tests display, as opposed to essay examinations:

Summary of Advantages of Objective Tests

over Essay Examinations

Objective Tests	Essay Examinations
1. Content validity is high. a) Adequate sampling of content. b) The task presented to the student is well defined. c) A multitude of the student's abilities are appraised.	1. Content validity is questionable. a) Inadequate and narrow sampling. b) The task presented to the student is vague and undefined. c) A very limited number of student's abilities are appraised.
2. Reliability is high. a) Responses are limited. b) Error of measurement due to the test content is very low. c) Scoring is very reliable: scorers' errors are eliminated; no error due to inconsistency among different scorers; no error due to inconsistency within the same scorer. d) Free from factors of skill and penmanship.	2. Reliability is usually very low. a) Responses are not limited. b) Error of measurement due to content is high because of the subjectivity and narrow sampling of the questions. c) Scoring is unreliable: two human errors exist in scores: the error due to the inconsistency among scorers, and the error due to inconsistency within the same scorer. d) Factors of skill and penmanship are uncontrolled.
3. Provides useful data for the interpretation of students' attainment. a) Results are easily interpreted by using objective procedures. b) Provide norms which describe average or typical performance, which can be compared with that of previous years.	3. Data provided cannot be relied upon for interpreting student's attainment. a) No objective procedures can be applied to interpret results. b) Typical or average performance is difficult to identify, and cannot be compared with typical performance of other years.
4. Standardized forms can be produced to serve various purposes.	4. Essay tests are used for limited purposes.

Objective Tests	Essay Examinations
5. Study habits of students are such that they pay attention to the details as well as to the generalities of the subject.	5. Students read generalities without paying much attention to the details.
6. Writing a response (by the student) is brief and requires a minimum of time.	6. Writing a response is time-consuming.
7. Each item has a predetermined key. Scoring can be done quickly, by a clerk; a specialist is not needed.	7. A standard key cannot be prepared. Any model answer is modifiable. The rating of papers is a tedious lengthy and inaccurate process. Rating must be done by a specialist.
8. Expenditure of time and money is minimum. Cost and time of both administration and scoring are relatively very small.	8. Time , money, and energy-consuming. Cost and time of both administration and scoring are considerable.

C H A P T E R 5

P R O P O S A L S

As a result of the study in this thesis two proposals are made: (A) If essay examinations are to be maintained they can and should be considerably improved. (B) Preliminary steps should be taken for using standardized objective tests for a substantial part of the educational measurement required.

A. Suggestions for Improving Essay Examinations:

Many of the general principles recommended for constructing objective tests apply to essay examinations as well. These principles and others are considered below.

I. Planning the Questions:

A plan must be thoughtfully worked out before questions are written. Such a plan would include the following:

(1) The function of the examination: This must be clearly defined by consulting the official statement of the educational policy and by consulting with those persons whose positions entitle them to define the policy of examinations. It may be necessary to state the function in specific terms. Thus if it is desired to test the ability of students to memorize material prescribed in the textbook, this must be explicitly stated.

(2) The course-objectives that will be included in the questions: The sources useful in formulating a careful statement of objectives are the syllabus, books on teaching the course, subject specialists, and experienced teachers. Objectives identified may then be classified into major types, and each type defined in terms of course content and expected student behavior.

(3) Sources: It is important to define the course content by proper choice of textbooks and other sources of material in which the scope and degree of complexity of each part is most appropriate for the level of the examination for which questions are to be set.

(4) Outline of course-content: After carefully studying the course content in the sources obtained, an outline of the content is prepared. In this outline all the facts, and concepts belonging to a certain topic are listed under it. The attempt should be made to have all the topics of the syllabus represented in the questions. The number of questions to be set on each topic can be approximately determined by what emphasis is given to one topic relative to the others. Emphasis is best judged by the proportion of class and assignment time which it is expected a teacher will devote to that topic. Then, when it is known how many questions can be given during one examination, the corresponding proportions of questions on time should be assigned to each topic.

II. Writing the Questions:

Before starting to write questions, the writer is advised to study the different ways of stating a question and to choose those that best suit his purpose. Generally, the type of structure chosen for the statement of a question should be appropriate to the level of knowledge to be tested, to the type of objective to be included, and to the mental process to be brought out.

The following rules for writing essay questions have been suggested by authorities on testing:

(1) "No question should be written until its purpose has been clearly defined."[1]

(2) "Write the question in such a way that the task is clearly and unambiguously defined for each examinee."[2]

(3) "In general start a question with such phrases as "Compare", "Contrast", "Give the reasons for", "Present the argument for and against", "Give original examples of", and "Explain how or why". Don't start essay questions with such words as What, Who? When, and List. These words are likely to present tasks requiring only the reproduction of information."[2]

---

[1] C. C. Ross, Measurement in Today's Schools, p. 164, Prentice Hall, Inc., 2nd edition 1953.

[2] R. L. Thorndike and E. Hagen, Measurement and Evaluation in Psychology and Education, pp. 53-55, New York, London, John Willey & Sons Inc. 2nd. edition, 1961.

(4) "The words What do you think, In your opinion, or Write all you know about almost never belong in an essay question to measure academic achievement."[1]

(5) "There is some evidence that a more valid sampling of the pupil's knowledge is afforded by increasing the number of questions and reducing the length of discussion expected on each. In many cases a well-constructed paragraph is sufficient . . . In any case the question should be so worded as to restrict the responses toward the objective which it is desired to measure."[2]

(6) " Be sure that the students do not have too many or too lengthy questions to answer in the time available."[1]

(7) "Have each examinee answer the same questions. Do not offer a choice of questions to be answered. Giving a choice of questions reduces the common base upon which different individuals may be compared. It adds one further source of variability to the subjectivity and inaccuracy that already exist."[1]

III. Scoring:

(1) The writer of the questions should prepare model answers showing what points are to be covered and how many credits

---

[1] Ibid.

[2] C. C. Ross, op. cit., p. 165.

are to be allowed for each. The model answer should specify exactly how many score points, if any,such features as organization, handwriting, sentence structure, etc. should be awarded.

2) The committee on scoring the answers should revise and check the model answers prepared by the writers of the questions. Several papers of student answers are read and checked against the model answer. Changes in the model answer and in its scoring resulting from the survey of student answers must be made before assigning scores to papers.

3) The answers to the same question are to be scored by a committee of at least two persons; each doing a thorough independent reading and rating, and without knowledge of other's rating. It is better that one scores the papers after they have been arranged in a different order.

The ratings on each paper are compared; where differences occur, the scorers re-read the answers together and arrive at an agreement if possible. Otherwise further persons are consulted, or, if the differences between the ratings are not grave, the average is taken.

B. Preliminaries for Using Objective Tests

I. As a first step, standardized tests can be used in the Preparatory Examination. For this purpose, objective tests must be prepared and tried out in a manner somewhat similar to that described in chapter 3, after making necessary refinements. The refinements are

to be made in the construction of the tests and in arranging for the first trial. In the construction of the tests, extensive use has to be made of various resources, of subject-specialists, and of teachers of preparatory classes. For the first trial of a locally-prepared test, a random sample of 250 students is a reasonable sample size. If the test passes the first trial and is then administered to the examinees in the General Preparatory Examination, the second administration will furnish the data necessary for standardization.

II. The first use of objective tests in public examinations might be attended by certain difficulties. The students are not accustomed to the technique. Teachers in general are not very well acquainted with objective tests. There are those who will claim that objective tests are unsatisfactory for testing certain important functions. Also it is to be expected that certain segments of public opinion will be agitated by any change in the method of testing, unless the change is properly introduced by presenting the merits of the new method.

To obtain acceptance, appreciation, and understanding of the idea of objective tests the following steps are recommended:

(1) Teachers are urged to use objective tests in their classes. Actually, the Ministry of Education recently distributed a circular to teachers advising them to use objective tests in school examinations. The circular did not have the desired effect, since not

many teachers know how to construct and use objective tests. Some of the ways of instructing teachers on the use of objective tests are: (a) through the visits of inspectors, (b) through summar and in-service training programs, (c) by making available to teachers books on tests and testing, (d) by including in the circulars to teachers general directions for the construction and use of objective tests, and (e) by distributing to teachers sample objective tests in their subject fields.

(2) Objective tests can be one of the effective methods used by administrators and supervisors for evaluating the curriculum, the effectiveness of instruction, or the educational program as a whole. A test that is designed to test scholastic achievement of students can give useful information about the efficiency of teaching and the effectiveness of a particular area in the curriculum.

(3) The provision of persons specialized in the construction, use, and interpretation of results of tests is a pre requisite for any success in introducing the idea of objective tests. This can be accomplished by offering a number of scholarships for full specialization in a university or for a short period of training in a recognized institute for testing. It would be very useful to have a special section in the Ministry of Education concerned with testing. If such a section is properly equipped with qualified persons, it would, among other things, help in training persons needed for subsequent expansion of testing programs.

III. Public examinations are but one method of evaluating the educational product. Objective tests are recommended in so far as they contribute to the accuracy of data provided by the examinations. Whatever the accuracy may be, theorists may claim that the evaluation based on the data of the examinations alone is incomplete, since such data constitute only partial information about the student. However since the best evaluation can only be based on such accurate information as is available, evaluation of the student in a public examination that has definite purposes will be confined to the examination results, made as accurate as possible, untill other procedures are developed that make evaluation more comprehensive.

The term "testing programs" refers to a planned procedure or procedures for the purpose of evaluating certain outcomes. In education, testing programs are concerned with the evaluation of the educational product, the educational program, or specific problems in administration, instruction, and the curriculum.

Testing programs can then be started as soon as all necessary provisions and facilities become available.

One of the advantages of testing programs will be mentioned. In these programs standardized tests are used. One of the important characteristics of these tests is that they furnish norms by which the scores are interpreted.

A norm is usually the average or typical value of a particular characteristic measured in a specified group such as the population of pupils at a certain age or in a certain school grade.

In most cases, the norm is the mean or median, but sometimes other points such as percentile or standard scores are used. However, since the norm is derived from the raw scores on the test, the usual practice is to prepare a table in which more than one type of score equivalent is shown, so as to provide a clear and detailed picture that will be useful for further application and interpretation of the test results. In any case, the norm is a reference point presumed to represent the level of attainment of the specified group.

Norms are developed in the process of standardizing tests. These tests may be achievement tests, aptitude tests, intelligence tests, or any other test designed for measuring one or more of the psychological characteristics of the individual.

In Jordan, no attempt has been made to establish local norms, or to compare, when comparison is possible and useful, with other already established norms. Testing programs, therefore, should aim at establishing local norms, and, whenever it is possible, comparing local attainment with outside norms.

Such a task is not an easy one. One of the difficulties appears to be in developing, or adapting, intelligence tests. A testing program designed for measuring intelligence would require such an expenditure of money and technical skill that it is questionable that the Ministry of Education in Jordan would, alone, undertake such a task. It seems that a great success can be accomplished if

intelligence tests are prepared for Arab children in general and not only for Jordanians. Some Arab countries, then, possibly with the assistance of cultural agencies such as U.N.E.S.C.O. would cooperate to undertake such a useful project.

On the other hand, local norms of achievement or aptitude can be gradually developed, to the extent that available means permit. The Ministry of Education is the agency best fitted to initiate and carry out testing programs. The extent to which the Ministry recognizes the need for and the importance of these programs will, to a large extent, determine how much progress will be made.

Uses of Norms:

At this point, it is useful to present some of the uses of norms which would have immediate practical applications to educational needs in Jordan:

(1) Norms provide bases for interpreting the scores of an individual by relating them to the scores of the group (the normative group). For example, if a standard test, say on arithmatic, is designed for the sixth elementary grade, then it is possible to know what score the average boy has obtained in this test over, say, the past three years; what percentage of the total group of sixth graders lies above or below a certain score; how well high, medium, and low scorers succeeded in future academic work (i.e., how well the scores may be used to predict success in future work).

(2) Norms can be used for prediction of future success. This implies that the standardized test for which norms have been derived should have predictive validity. Predictive validity can be ascertained by correlating the test scores with future achievement, when data on the latter can be easily secured. For pupils proceeding from one class to another within the educational ladder, the school grades acquired at a certain class or stage will serve as criteria for validating a test designed for a lower class or stage. For pupils who complete secondary education, prediction studies become rather complicated because of the difficulty of getting data about their future work. However, validation studies in this case are important for two reasons, they are useful for (a) improving the educational program of secondary schools, and (b) improving the qualities of tests designed for pupils in the secondary stage. The data needed for prediction studies may be obtained in two ways, (a) by collecting university grades of graduates who go to universities, and (b) through follow up studies on work records, promotions, recommendations of supervisors or employees, etc. of those graduates who do not go to universities but are employed in various kinds of jobs.

(3) Norms provide bases for comparing the attainment of a particular group in a certain area or school with the attainment of the normative group. The comparison is done between the average performance of the particular group on a test or group of tests

with that of the normative group. Such a comparison is useful for discovering weaknesses that should be remedied.

Similarly, two schools can be compared with each other, a private school with a public one, a school in a rural area with a school in an urban area, etc.

(4) Norms are useful for diagnostic purposes. The class-room teacher who uses standardized tests will be able to understand more about his students by comparing their achievement on the test with the specified norm.

The effectiveness of the curriculum in developing the students' knowledge and abilities may be studied by administering standardized tests to all, or to a selected sample, of the students learning the curriculum. The scores obtained on the tests would show where the curriculum was successful or unsuccessful in producing effective learning.

(5) Norms are useful for placement and guidance purposes. A boy who scores low on the test, relative to the group, needs special attention. On the other hand, the boy who scores high may be given extra work that would satisfy his talents. A group of tests (aptitude and achievement) administered to students will yield for every student a number of scores by means of which a correct judgement is more likely to be made as to the kind of learning most suitable to him. Thus students chosen for vocational schools would be selected

on the basis of their demonstrated abilities and aptitudes to profit from such training.

(6) Norms are useful for educational research studies:

a) A certain program is tried in a school. The effectiveness of the program is judged by matching the attainment of students in this school on a test or tests with that of students in other schools.

b) A method of teaching is tried with one class. In another class, another method is used. Other variables, such as age, intelligence, background, etc., in both classes are controlled. The two methods of teaching are compared by comparing the performance of both classes on the same test, for which a norm is available.

c) A change in the curriculum content is proposed. The effect of the change is studied by comparing the scores of students with whom the changed curriculum was tried on a test with the scores of other students that have been taught under the unchanged curriculum.

(7) Norms are useful in studying the development and progress of education over a number of years. Thus the performance of students on a test or tests in 1965 is compared with the performance on the same test in 1960 or in every successive year. Any improvement in average performance on the test will show that absolute

progress has taken place in the area of study on which the test is constructed.

It must be pointed out that norms do not remain the same as time passes on. They usually change as the factors that make them up change. But they can be very useful elements for guiding the change toward improvement and progress.

A P P E N D I X I

Experimental Objective Achievement Test

in

Mathematics

اختبار تحصيل موضوعي
في الرياضيات
(مستوى امتحان الشهادة الاعدادية العامة)
للمدارس الاردنية

الزمن : خمسون دقيقة

تعليمات

١ـ ارجو ان لا تكتب شيئا على هذا الكراس ، يجب ان تظهر اجاباتك على "ورقة الاجابة" فقط. استعمل الورقة المنفصلة الموجودة داخل الكراس للمسودة ، سلم هذه الورقة مع الكراس وورقة الاجابة عند انتهاء الاختبار

٢ـ لا تقلب هذه الصفحة الا عندما يُعلن عن ذلك

٣ـ يتكوّن هذا الاختبار من عدد من الاسئلة اعطي لكل منها اربع اجابات احداها فقط هو الصحيح . اقرأ السؤال والاجابات المذكورة بعده ، وبعد ان تختار الجواب الصحيح عيّن رقم السؤال على "ورقة الاجابة" واصنع بالقلم خطا مائلاً هكذا / يقطع الحرف المماثل للحرف الذى يسبق الجواب الذى اخترته .

مثال ١ مجموع زاويا المثلث بالدرجات الستينية يساوي

أ ٣٦٠°

ب ٢٧٠°

جـ ١٨٠°

د ٩٠°

انظر الى ورقة الاجابة عند مثال ١ ولاحظ ان الاشارة / قد رُسمت قاطعة الحرف جـ وهو نفسه يسبق الجواب ١٨٠° في المثال ١.

٤ـ عددُ الاسئلة في هذا الاختبار ثلاثون ، والوقت المحدد لاجابتها جميعا هو خمسون دقيقة اى ان معدل الزمن الذى يحتاجه السؤال الواحد يزيد قليلا على دقيقة ونصف

٥ـ علامة الاجابة الكاملة ستون ، اى ان لكل سؤال علامتين ، وتحسب علامتك بعد خصم جزء من اجاباتك الخاطئة من مجموع اجاباتك الصحيحة ولذلك فليس من صالحك ان تخمّن الجواب الصحيح الا اذا استطعت ان تضيّق مجال الاختيار بين الإجابات المذكورة.

٦ـ من المحتمل ان لا يتمكن اىٌّ منكم من الاجابة على جميع الاسئلة ، حاول جهدك ، واذا بدا لك ان احد الاسئلة يتطلب وقتا اتركه على ان تعود اليه فيما بعد اذا بقي لديك وقت.

لا تقلب هذه الصفحة الا عندما يعلن عن ذلك

١ـ أيّ المقادير التالية يساوي المقدار (ـ ٢ س) ؟

أ $٣ س^{٢}$

ب $-٦ س$

ج $٨ س^{٢}$

د $-٨ س^{٣}$

٢ـ ما قيمة المقدار $\sqrt{\frac{١٨١ أ^{٢}}{٤ س^{٤}}}$ عندما تكون قيمته س = ٣ ؟

أ $\pm \frac{أ}{٢}$

ب $\pm ٢ أ$

ج $\pm \frac{٢ أ}{٢}$

د $\pm \frac{٣ أ}{٢}$

٣ـ يُعبّر عن الطاقة الكهربائية المتولدة في سلك يمر فيه تيار كهربائي بالمعادلة (ق = $ت^{٢}$ م ن) ، ما هي قيمة ت بدلالة الرموز الاخرى؟

أ $\frac{ق}{م ن}$

ب $\frac{ق م}{ن}$

ج $\sqrt{\frac{ق}{م ن}}$

د $\sqrt{ق م ن}$

٤ـ أحد جذري المعادلة ($س^{٢}$ = ٩س) هو

أ ٣

ب ـ ٣

ج صفر

د ـ ٩

٥ـ أيّ المقادير التالية يساوي المقدار ٢أ (س ـ ٥) (س + ٥) ؟

أ $٢ أ س^{٢} - ٢٥$

ب $٢ أ س^{٢} + ٥٠$

ج $٢ س^{٢} - ٢٥ أ$

٦ـ ما ناتج قسمة المقدار (أ + $\frac{١}{٨ س^{٣}}$) على المقدار (أ + $\frac{١}{٢ س}$) وذلك عندما تكون قيمة س = ٢ ؟

أ $\frac{٥١}{١٦}$

ب ٢

ج $\frac{١٣}{١٦}$

د $\frac{١٦}{١٣}$

٧ـ ما هي القيمة السالبة التي تنتج من الجذر التربيعي للمقدار ($س^{٤}$ ـ ٢ $س^{٢}$ + ١) ؟

أ ١ ـ س

ب ـ $س^{٢}$ ـ ١

ج ٢ ـ س

د ١ ـ $س^{٢}$

٨ـ ما هي المعادلة التي جذراها ـ ٥ ، ٥ ؟

أ $س^{٢} - ٥ س = ٠$

ب $س^{٢} - ٢٥ = ٠$

ج $(س - ٥)^{٢} = ٠$

د $س^{٢} - ٢٥ س = ٠$

٩ـ في المعادلة $\frac{٣}{م} + \frac{٢}{ص} = \frac{٧}{٢ س}$ ما هي قيمة س بدلالة م ؟

أ $\frac{٢}{م}$

ب $\frac{م ٥}{٦}$

ج $\frac{م ٥}{٣}$

د $\frac{م}{٣}$

١٠ـ ما قيمة س في المعادلتين الآتيتين التاليتين:
(154) (٢س ـ ص = ٤أ)، (س + ٣ص = ١٩أ) ؟

أ ٣أ
ب ٢أ
ج ـ٣أ
د أ

١١ـ اشترى شخص أ مترا من قماش بسعر د فلسا للمتر الواحد وباعها جميعا بمبلغ ك فلسا فكم فلسا ربح في المتر الواحد؟

أ $\frac{ك}{أ}$ ـ د

ب $\frac{ك}{أ}$ + د

ج ك ـ د

د ك ـ أد

١٢ـ ما قيمة ص في المعادلتين التاليتين
(س ص = ٣٦)، (ص + ٣) (س ـ ٨) = ٢٠

أ ١٢
ب ٨
ج ٦
د ٣

١٣ـ اذا كان الخط البياني للمعادلة (٣س + ٢م ص = ٩) يمر بالنقطة (٣، ١) فما هي قيمة م؟

أ ٢
ب ٦
ج ٤
د ٣

١٤ـ أي القيم التالية يساوي المقدار $\frac{٦ب س ـ ٤ب^{٢} ـ ٢س^{٢}}{س ـ ٢ب}$

أ ٢ب ـ س
ب ٢(ب ـ س)
ج ٢س ـ ٢ب
د س ـ ب

١٥ـ أي القيم التالية يعبر عن المبلغ الذي تصبح جملته ج دينارا اذا وضع في بنك ليربح ربحا بسيطا بسعر د ٪ لمدة ن سنة؟

أ $ح + \frac{ح د ن}{١٠٠}$

ب $\frac{١٠٠ ح}{١٠٠ + د ن}$

ج $\frac{ح د ن}{١٠٠}$

د $\frac{١٠٠ ح}{د ن}$

١٦ـ أمّن شخص على حياته بمبلغ ٥٠٠ دينار على ان يدفع سنويا ١٠٪ من المبلغ، فاذا توفي بعد ٣ سنوات ما خسارة شركة التأمين لورثته؟

أ ٤٠٠ دينارا
ب ٣٥٠ دينارا
ج ٢٥٠ دينارا
د ١٥٠ دينارا

١٧ـ باع شخص سيارته بمبلغ ٧٥٠ دينارا، وبذلك خسر ٤٠٪ من ثمن شرائها، بكم اشترى السيارة؟

أ ١٠٠٠ دينار
ب ١٢٠٠ دينار
ج ١٢٥٠ دينار
د ١٥٠٠ دينار

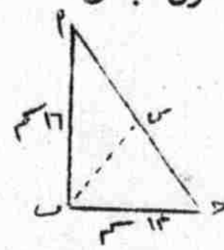
١٨ـ اشترى مصنع آلة بمبلغ ١٠٠٠ دينار، فاذا تناقصت قيمة الآلة نتيجة الاستعمال بمعدل ١٠٪ من قيمتها في السنة السابقة، كم تصبح قيمتها بعد سنتين؟

أ ٨١٠ دينار
ب ٨٠٠ دينار
ج ٧٩٠ دينار
د ١٢١٠ دينار

١٩ـ اشترى رجل ١٠٠ سهم بسعر السهم ٢٠ دينارا وثمنه الاساسي ١٠ دنانير، فاذا وزعت الشركة ارباحها بمعدل ١٥٪ في السنة، ما هو دخله السنوي من اسهمه؟

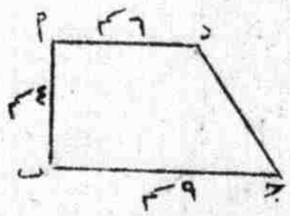
أ ١٥٠ دينارا
ب ٣٠٠ دينارا
جـ ١١٥٠ دينارا
د ٢٣٠٠ دينارا

٢٠ـ الشكل المجاور يُمثِّل مثلثاً قائم الزاوية في ب، النقطة س تنصف الوتر أ جـ، ما هو طول ب س؟



أ ٢٠ سم
ب ١٢ سم
جـ ١٠ سم
د ٧،١ سم

٢١ـ الشكل المجاور يُمثِّل شبه منحرف قائم الزاوية في ب، ما هو طول د حـ؟



أ ٥ سم
ب ٦ سم
جـ ٧ سم
د ٣ سم

٢٢ـ أيُّ العبارات الاربع يكمل الجملة التالية فيُكسبها مفهوما ينطبق في جميع الاحوال، المستقيم الواصل من راس المثلث الى منتصف القاعدة ...

أ ينصف زاوية الراس
ب يقسمه الى مثلثين منطبقين
جـ يكون عموداً على القاعدة
د يقسمه الى مثلثين متكافئين

٢٣ـ معين طول ضلعه ١٠ سم وطول احد قطريه ٢٠ سم، ما هو طول القطر الآخر؟

أ ٨،١٢٥ سم
ب ١٢ سم
جـ ١٠ سم
د ٦ سم

٢٤ـ اذا كان ارتفاع مثلث متساوى الاضلاع $٨\sqrt{٣}$ سم، ما هو طول ضلعه؟

أ ٣٢ سم
ب ٢٤ سم
جـ ١٦ سم
د $١٦\sqrt{٣}$ سم

٢٥ـ وتر في دائرة طوله ٦ سم وطول العمود النازل عليه من المركز ٤ سم، ما هو نصف قطر الدائرة؟

أ ٨ سم
ب ٧،٢ سم
جـ ٥ سم
د ٤ سم

٢٦ـ أ ب جـ أيُّ مثلث، أ د يُنصِّف الزاوية أ ويقطع ب جـ في د، أيٌّ من العبارات التالية ينطبق في جميع الاحوال؟

أ أ د عمود على ب جـ
ب أ د يُنصِّف مساحة المثلث أ ب جـ
جـ نقطة د تُنصِّف ب جـ
د نقطة د على بعدين متساويين من أ ب، أ جـ

٢٧ـ احدى العبارات التالية تنطبق في جميع الاحوال

أ اذا تقاطع وتران متساويان داخل دائرة نصَّف كلٌّ منهما الآخر.
ب اذا اختلف طولا وترين في دائرة كان اصغرُ الوترين أقربهما الى المركز.
جـ اذا تساوى طول وتر في دائرة مع طول وتر في دائرة اخرى كان قطرا الدئرتين متساويين
د اذا تساوى بعدا وترين في دائرة عن مركزها كان الوتران متساويين

٢٨ـ احدى العبارات التالية تنطبق في جميع الاحوال :

أ نقطة التقاء المستقيمات المتوسطة في المثلث تكون على ابعاد متساوية من اضلاعه

ب نقطة التقاء منصّفات زوايا المثلث تكون على ابعاد متساوية من اضلاعه

جـ نقطة التقاء ارتفاعات المثلث تكون على ابعاد متساوية من اضلاعه

د نقطة التقاء الاعمدة المنصّفة لاضلاع المثلث تكون على ابعاد متساوية من اضلاعه

٢٩ـ أي من النقاط التالية يمكن ان تكون مركزا لدائرة تمر برؤوس مثلث ؟

أ نقطة تقاطع المستقيمات المتوسطة للمثلث

ب نقطة تقاطع منصّفات زوايا المثلث

جـ نقطة تقاطع الاعمدة المنصفة لاضلاع المثلث

د نقطة تقاطع ارتفاعات المثلث

٣٠ـ احدى العبارات التالية خطأ لا تنطبق في جميع الاحوال

أ كل متوازى اضلاع دائريّ يجب ان تكون كلّ من زواياه قائمة

ب في كل شكل رباعي دائرى يتقاطع القطران في مركز الدائرة التي تمر حوله

جـ في كل شكل رباعي دائرى ، كل زاويتين متقابلتين تكونان متكاملتين

د كل قطرين في دائرة يمكن ان يكونا نفسيهما قطرى مستطيل في نفس الدائرة

(157) ٢٧

(14)

ثانوية الجميه

اختبار تحصيل موضوعي في الرياضيات

Answer Sheet

ورقة الاجابة

(46)

الاسم: فوزي محمد الصف: أول ثانوي التاريخ: ١١/٩/٢٠٠٢م

مثال ١ أ ب ج د

١	أ	ب	ج	د	١١	أ	ب	ج	د	٢١	أ	ب	ج	د
٢	أ	ب	ج	د	١٢	أ	ب	ج	د	٢٢	أ	ب	ج	د
٣	أ	ب	ج	د	١٣	أ	ب	ج	د	٢٣	أ	ب	ج	د
٤	أ	ب	ج	د	١٤	أ	ب	ج	د	٢٤	أ	ب	ج	د
٥	أ	ب	ج	د	١٥	أ	ب	ج	د	٢٥	أ	ب	ج	د
٦	أ	ب	ج	د	١٦	أ	ب	ج	د	٢٦	أ	ب	ج	د
٧	أ	ب	ج	د	١٧	أ	ب	ج	د	٢٧	أ	ب	ج	د
٨	أ	ب	ج	د	١٨	أ	ب	ج	د	٢٨	أ	ب	ج	د
٩	أ	ب	ج	د	١٩	أ	ب	ج	د	٢٩	أ	ب	ج	د
١٠	أ	ب	ج	د	٢٠	أ	ب	ج	د	٣٠	أ	ب	ج	د

$17 - \frac{10}{3}$

١٧ ٢٧

A P P E N D I X II

Experimental Objective Achievement Test

in

Sciences

اختبار تحصيل موضوعي
في العلوم
* مستوى امتحان الشهادة الاعدادية العامة *
للمدارس الاردنية

الزمن : خمسون دقيقة

تعليمات

١ـ الرجاء أن لا تكتب شيئاً على هذا الكراس ، يجب ان تظهر اجاباتك على * ورقة الاجابة * المرفقة فقط .

٢ـ لا تقلب هذه الصفحة الا عندما يُعلن عن ذلك

٣ـ لا تبدأ الاجابة الا بعد ان تقرأ جيّدا التعليمات المذكورة هنا وبعد ان تفهم طريقة الاجابة

٤ـ يشتمل هذا الاختبار على عدد من الاسئلة أُعطي لكل منها ثلاث او اربع اجابات احدها فقط هو الصحيح . بعد ان تقرأ السؤال والاجابات المذكورة بعده تبحث عن رقم السؤال على * ورقة الاجابة * وترسم خطا مائلا هكذا / يقطع الحرف المماثل للحرف الذى يسبق العبارة او الكلمة التي اخترتها لتكون الاجابة الصحيحة والمثال التالي يوضح ذلك :

مثال (١) وحدة قياس الطول المترية هي :

أ الغرام

ب السنتمتر

جـ الليتر

د الدقيقة

الاجابة الصحيحة هي * السنتمتر * انظر الى ورقة الاجابة عند (مثال ١) ولاحظ ان الاشارة / قد رسمت قاطعة الحرف ب وهو نفسه يسبق كلمة السنتمتر في هذا المثال .

٥ـ عدد الاسئلة في هذا الاختبار اربعون لكل منها علامة واحدة والوقت المحدّد لاجابتها خمسون دقيقة .

٦ـ تحسب علامتك على هذا الاختبار بعد خصم جزء من اجاباتك الخاطئة من مجموع اجاباتك الصحيحة ولذلك فليس من صالحك ان تخمّن الجواب الا اذا استطعت ان تضيق مجال اختيارك للجواب الصحيح .

٧ـ من المحتمل ان لا يتمكن اى منكم من الاجابة على جميع الاسئلة في الوقت المحدد ، حاول جهدك فاذا بدا لك انّ سؤالاً يتطلب وقتاً للتفكير اتركه على ان تعود اليه فيما بعد اذا تبقى لديك وقت .

لا تقلب الصفحة الاّ عندما يُعلن عن ذلك .

١ـ اذا كان الوزن الذري لعنصر ٢٣٥ وعدد النيوترونات في ذرته ١٤٦ ، فما هو العدد الذري للعنصر ؟
أ ٢٣٥
ب ١٨٩
جـ ١٤٣
د ٨٩

٢ـ تقول نظرية الحركة الجزيئية أنه عندما يتحول الجسم الصلب الى غاز
أ يزداد حجم جزئيات الجسم
ب يقلّ وزن جزئيات الجسم
جـ تزداد المسافات بين الجزيئات
د تزداد جاذبية التماسك بين الجزيئات

٣ـ اذا كان الوزن الذري للهيدروجين = ١ وللاكسجين = ١٦ وللنيتروجين = ١٤ يكون الوزن الجزيئي لهيدروكسيد الامونيوم
أ ٧٠
ب ٣٥
جـ ٣١
د ٥٣

٤ـ عندما يتفاعل اكسيد الخارصين مع حامض الكبريتيك فان احد نواتج التفاعل هو
أ الهيدروجين
ب الاكسجين
جـ كبريتات الخارصين
د اكسيد الكبريت

٥ـ يتكون الجبس (الجبصين) من
أ كبريتات الكالسيوم
ب كبريتات المغنيسيوم
جـ كربونات الصوديوم
د كربونات المغنيسيوم

٦ـ أهم مركبات الفوسفور المستخدمة في صنع الاسمدة الكيماوية هو
أ خامس اكسيد الفوسفور
ب فوسفات الكالسيوم
جـ حامض الفوسفوريك
د فوسفات المغنيسيوم

٧ـ أحد الخواص الكيماوية التالية ينطبق على حامض النيتريك
أ عامل مؤكسد قوي
ب عامل مختزل قوي
جـ يُطلق الهيدروجين بتفاعله مع الفلزات الثقيلة
د قليل الذوبان في الماء

٨ـ التفاعل الكيماوي الرئيسي في تحضير الحديد في الفرن النفاخ هو
أ اختزال اكاسيد الحديد
ب تحويل خامات الحديد الى اكاسيد
جـ تنقية الحديد من الشوائب
د تحويل خامات الحديد الى كربوناته

٩ـ يشمل الجدول المسمى " السلسلة الدافعة الكهربائية الكيمائية " على
أ العناصر مرتبة حسب اعدادها الذرية
ب العناصر مرتبة في عائلات لكل منها خواصها الكيماوية
جـ العناصر الفلزية حسب قدرة احدها ان يحل مكان الآخر في مركباته
د العناصر اللافلزية مرتبة حسب قدرتها على تكوين احماض.

١٠ـ احد الخواص الكيماوية التالية ينطبق على عنصر البوتاسيوم
أ اكسيده قلوي
ب لا يتفاعل مع الحوامض المخففة
جـ عامل مؤكسد
د مركباته قليلة الذوبان في الماء

١١ـ عندما تفاعلت ٣٫٦ غم من المغنيسيوم تفاعلاً تاماً مع حامض الكلوريد ريك اطلقت ٠٫٣ غم من الهيدروجين . اذن يكون الوزن المكافئ للمغنيسيوم
أ ٢٤
ب ١٨
جـ ١٢
د ٦

١٢ـ يتفاعل غرام واحد من الهيدروجين مع اربعة غرامات من الاكسجين فينتج عن ذلك ان
أ يتبقى ٠٫٧٥ غم من الهيدروجين
ب يتبقى ٢ غم من الاكسجين
جـ يتكون ٤٫٥ غم من الماء
د يتكون ٥ غم من الماء

١٣ـ وظيفة الشعيرات الجذرية في النبات هي
أ مساعدة الجذر على اختراق التربة
ب حماية الجذر من التآكل في التربة
جـ امتصاص الماء والغذاء
د اختزان الغذاء

١٤ـ جزء الساق الذي ينقل العصارة الغذائية من الجذر الى الاوراق هو
أ اللحاء
ب النخاع
جـ القشرة
د الخشب

١٥ـ الوظيفة النباتية التي بسببها لا يستحسن وضع نباتات في غرف النوم هي
أ التمثيل الضوئي
ب النتح
جـ امتصاص الاغذية
د التنفس

١٦ـ الزهرة الخنثى هي التي
أ تحوى اعضاء التأنيث والتذكير معا
ب تحوى عضو التأنيث أو عضو التذكير
جـ لا يتم فيها التلقيح
د لا يتم فيها الاخصاب

١٧ـ تسمى العملية الرئيسية التي يتم بواسطتها امتصاص النبات للمواد الغذائية من التربة
أ الخاصية الشعرية
ب الانتشار الغشائي
جـ التمثيل الكلوروفيلي
د الانتشار الميكانيكي

١٨ـ عضو التنفس الرئيسي عند صغار الضفادع هو
أ الرئة
ب الخيشوم
جـ الفم
د الجلد

١٩ـ وظيفة الشرايين في الارنب هي
أ نقل الدم النقي من القلب الى اجزاء الجسم
ب نقل الدم غير النقي من الجسم الى القلب
جـ نقل الدم النقي من الرئتين الى القلب

٢٠ـ وظيفة الحجاب الحاجز في الارنب هي
أ فصل التجويف الصدري عن تجويف البطن
ب تنظيم حركات المعدة اثناء الهضم
جـ تنظيم عملية التنفس
د تنظيم افراز الغدد الهاضمة

٢١ـ متى يتم اخصاب بيضة الدجاجة ؟
أ عندما ترقد الدجاجة الملقحة على البيضة
ب قبل خروج البيضة من الدجاجة الملقحة
جـ عندما تتوفر الحرارة والرطوبة الكافيتين في آلة تفقيس البيض
د يضع الدجاج ملقحا كان ام غير ملقح بيضا مخصبا .

٢٢ـ عندما يتحور عضو حيوي فانه
أ يصاب باعراض مرضية
ب يصبح اكثر ملائمة لوظائف اخرى
جـ يفقد خصائصه ووظائفه الاصلية
د يتوقف نموه

٢٣ـ في تجربة لمعرفة تأثير سماد معدني جديد على نوع من النبات ، زرعت نبتة في مكان مظلل بعد إضافة السماد الى تربته ، وزرعت نبتة أخرى في مكان آخر كثير الشمس ، وبعد اسبوع تبين أن النبته الاولى ذبلت وجفت بينما ظلت الثانية يانعة حية ، ماذا يمكن ان تستنتج من هذه التجربة ؟

أ يحتاج هذا النوع من النبات الى شمس كثيرة

ب السماد المعدني الجديد يضر هذا النوع من النبات

ج هذا النوع من النبات يحتاج الى السماد المعدني الجديد وكذلك الى الشمس الكثيرة

د لا يوجد ما يبرر اياً من النتائج أ أو ب أو ج

٢٤ـ حسب نظرية الالكترونات ، يكتسب جسم ما شحنة موجبة عندما

أ يفقد جزءا من بروتوناته

ب يفقد جزءا من الكتروناته

ج يكتسب جزءا من الكترونات جسم آخر

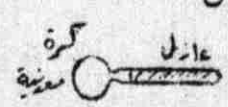
٢٥ـ عند تقريب قضيب موجب التكهرب الى قرص كشاف كهربائي غير مشحون ـ دون ان يلامس القضيب القرص ـ فان ورقتي الكشاف تنفرجان لأنّ

أ شحنةً تكوّنت بالتأثير على اجزاء الكشاف

ب جزءاً من شحنة القضيب انتقل الى القرص والورقتين

ج شحنةً انتقلت من الورقتين الى القضيب

٢٦ـ كرة معدنية مشحونة ، امسك بها شخص من مقبضها العازل وجعلها تلامس الجدار الداخلي لكأس معدني معزول ثم اخرجها وقربها الى قرص كشاف كهربائي فلم تنفرج الورقتان . اذن تكون الكرة فقدت شحنتها لان هذه الشحنة تسربت الى



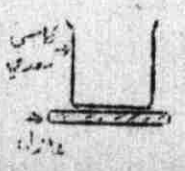
أ يد الشخص

ب السطح الداخلي للكاس

ج السطح الخارجي للكاس

د جميع اجزاء الكاس

هـ الهواء المحيط



٢٧ـ ينتج عن الاستقطاب في العمود البسيط :

أ تقوية التيار المتولد في الدائرة الخارجية

ب تقوية القوة الدافعة عند قطبي العمود

ج زيادة المقاومة الداخلية للعمود

د تفاعل النحاس مع حامض الكبريتيك

٢٨ـ عمل المستقبل التلغرافي مبنيٌّ على خواص :

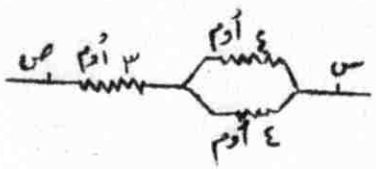
أ رموز مورس

ب التفريغ الكهربائي

ج التيار التأثيري

د المغناطيس الكهربائي

٢٩ـ في الشكل المجاور الذي يمثّل جزءا من دائرة كهربائية ، المقاومة المحصورة بين س ، ص هي :



أ ٣٫٥ اوم

ب ٥ اوم

ج ٧ اوم

د ١١ اوم

٣٠ـ اذا استعمل "العاكس" في مولد كهربائي تكون وظيفته :

أ نقل تيار ثابت الاتجاه الى الدائرة الخارجية

ب مضاعفة القوة الدافعة المتولدة في ملفات المولد

ج جعل التيار في ملفات المولد ثابت الاتجاه

د تنظيم حركة الدرع

٣١ـ المبدأ الاساسي في التاثير المغناطيسي للتيار هو

أ يتولد تيار كهربائي في موصل يقطع مجالا مغناطيسيا

ب يتولد مجال مغناطيسي حول موصل يحمل تيارا

ج يتمغنط قضيب من الحديد بالتاثير عند تقريبه من مغناطيس آخر

د تتولد حرارة في سلك يحمل تيارا كهربائيا

٣٢ـ مرر تيار شدته (٢) امبير في محلول كبريتات النحاس مدة (١٠٠٠) ثانية فاذا كان المكافئ الكهربائي الكيماوي للنحاس ٠٫٠٠٠٣٣ تكون كتلة النحاس المترسبة :

أ ٠٫٣٣ غم

ب ٠٫٦٦ غم

ج ٣٫٣ غم

د ٦٫٦ غم

٣٣ـ مركز ثقل صفيحة غير منتظمة الشكل هو :-
أ الخط الذى تقع عليه نقطة تتزن فيها الصفيحة
ب الخط الذى يمر في وسط الصفيحة
جـ نقطة على الخط الذى يمر في وسط الصفيحة
د نقطة يمر بها خط عمل وزن الصفيحة

٣٤ـ اثّرت قوتان متساويتان بينهما زاوية قائمة على نقطة في جسم ، فاذا كان مقدار كل منهما (ق) يكون مقدار محصلتهما :
أ ٢ ق
ب ٢√ق
جـ ق√٢
د صفر

٣٥ـ إحدى العبارات التالية يتفق والمفهوم الحديث لقانوني حفظ الطاقة والكتلة :
أ كمية الطاقة في الكون ثابته دائما
ب كمية الكتلة في الكون ثابتة دائما
جـ مجموع كميتي الطاقة والكتلة ثابت دائما
د كل من كمية الطاقة وكمية الكتلة ثابت لا يتغير

٣٦ـ سقط جسم من السكون عن قمة بناء فوصل الارض بعد ثانية واحدة فاذا كان تسارع الجاذبية ٣٢ قدم / ث/ ث يكون ارتفاع البناء
أ ٦٤ قدما
ب ٤٨ قدما
جـ ٣٢ قدما
د ١٦ قدما

٣٧ـ اثرت قوة مقدارها (ق) على جسم كتلته (ك) وبعد ان قطع مسافة (ف) اصبحت سرعته (س) اذن يكون الشغل الناتج
أ ك س
ب ق س
جـ ك ف
د ق ف

٣٨/ ٤٠ قابل الاجهزة المكتوبة في العمود الاول مع انواع تحوّل الطاقة في العمود الثاني أشّر على ورقة الاجابة على الحرف الذى يدلّ على نوع تحوّل الطاقة لكل من الاجهزة المرقمة ٣٨ ، ٣٩ ، ٤٠ : اجابة المثال ٢ تظهر على ورقة الاجابة

العمود الاول	العمود الثاني
مثال؟ المروحة الكهربائية	أ من طاقة كهربائية الى طاقة حرارية
٣٨ المولّد الكهربائي (الدينامو)	ب من طاقة كيماوية الى طاقة كهربائية
٣٩ المكواة الكهربائية	جـ من طاقة ميكانيكية الى طاقة كهربائية

اختبار تحصيل موضوعي
في العلوم

60

40
28 / 12/3 = 24
28 - 12

Answer Sheet

ورقة الاجابة

الاسم : سعد عبد أبو حمادة

الصف : الثاني الأدبي

التاريخ : ١٢ / ٩ / ١٩٦٥م

(مثال ١) ا ب ج د

١ ـ ا ب ج د
٢ ـ ا ب ج د
٣ ـ ا ب ج د
٤ ـ ا ب ج د
٥ ـ ا ب ج د
٦ ـ ا ب ج د
٧ ـ ا ب ج د
٨ ـ ا ب ج د
٩ ـ ا ب ج د
١٠ ـ ا ب ج د
١١ ـ ا ب ج د
١٢ ـ ا ب ج د
١٣ ـ ا ب ج د
١٤ ـ ا ب ج د
١٥ ـ ا ب ج د
١٦ ـ ا ب ج د
١٧ ـ ا ب ج د
١٨ ـ ا ب ج د
١٩ ـ ا ب ج
٢٠ ـ ا ب ج د
٢١ ـ ا ب ج د
٢٢ ـ ا ب ج د
٢٣ ـ ا ب ج د
٢٤ ـ ا ب ج
٢٥ ـ ا ب ج
٢٦ ـ ا ب ج د
٢٧ ـ ا ب ج د
٢٨ ـ ا ب ج د
٢٩ ـ ا ب ج د
٣٠ ـ ا ب ج د
٣١ ـ ا ب ج د
٣٢ ـ ا ب ج د
٣٣ ـ ا ب ج د
٣٤ ـ ا ب ج د
٣٥ ـ ا ب ج د
٣٦ ـ ا ب ج د
٣٧ ـ ا ب ج د
٣٨ ـ ا ب ج د
٣٩ ـ ا ب ج د
٤٠ ـ ا ب ج د

(مثال ٢) ا ب ج د

A P P E N D I X III

Experimental Objective Achievement Test

in

English Language

8/3

(166)

17

ACHIEVEMENT TEST

OF ENGLISH LANGUAGE

(The Gen. Prep. Exam. Level)
(in Jordanian schools)

First Name: _____

~~Family Name~~: _____

Class: _____ Date: _____

Time: 50 minutes

Directions

(Do not turn this page until you are told to do so.)

This test contains several types of questions. To every question you are given three or four answers, but only one of them is the correct answer. You are required to find the correct answer and to make a circle around the letter in front of it. Directions are given at the beginning of each set of questions. Some examples are done for you to show you how to answer the questions.

Your mark on this test will be computed by subtracting a part of your wrong answers from your correct answers. It is therefore not to your advantage to make a guess unless you are sure that one or more of the choices is incorrect.

You may not be able to answer all of the questions in the time allowed, but do your best to answer as many as you can. If a question seems difficult, leave it and go on; you may return to it later if you have time.

(DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO DO SO.)

(167)

Directions: Make a circle around the letter in front of the word that fits best in the space of each of the following sentences.

Example:

A here
B hair
(C) hear

I don't _____ your voice.

(Note that the circle is drawn around C; which means that "hear" is the correct word that fits best in the space. Note also that you must not write anything in the space. Now answer the following questions yourself.)

1. ✓
A their
B they
(C) them

My friends asked me to go with _____ to the football match.

2. ✓
A is
(B) are
C was

Both of my books _____ useful.

3. ✓
(A) whom
B whose
C who

The boy with _____ I was talking is my classmate.

4. ✓
(A) know
B no
C now

Does he _____ the way?

5. ✓
A what
B which
(C) how

Show me _____ I must solve this problem.

6. +
(A) There
B Their
C Theirs

_____ house is large.

7. ✓
A as soon as
B when
(C) where

I placed my book _____ it could be seen.

8. +
A was built
(B) was building
C built

A new house _____ near our school.

9. ✓
A to leave not
(B) not to leave
C to not leave

The mother told her child _____ the door open.

10. ✓
(A) too
B to
C two

The day was _____ hot.

11. +
A hurt
B hurted
(C) have hurted

This is a dangerous game because I _____ myself in it.

12. ✓
A Don't
(B) Doesn't
C Do

_____ he ever play football?

9 - 3½ = 7½

13. +
A was writtem
B was writing
(C) had written
D wrote

This book _____ before the war.

14. A wrote
B write
C to write
D writing

He made me _____ the exercise.

15. A Because
B If
C Although
D So that

_____ John worked hard, he failed in the examination.

16. A speak
B to speak
C speaking
D spoken

He can _____ English.

17. A have you?
B had you?
C haven't you done?
D didn't you?

You haven't done the exercise, _____

18. A do I
B I did study
C did I
D I studied

Sami studied History and so _____.

19. A take
B taking
C to have taken
D to take

I used _____ my medicine every morning.

20. A break
B breaked
C broken
D broke

He has _____ the window.

21. A This
B That
C Those
D which

_____ book are you going to read?

22. A does he
B did he write
C didn't he
D had he

He wrote his composition, _____ ?

23. A can't drive
B cam drive
C can
D cam't

This man can't drive a car, can he? No, he _____.

24. A will solve
B would solve
C have solved
D are solving

If you were clever you _____ the problem.

Directions: Make a circle around the letter in front of the word (or words) which gives the best meaning of the expression in brackets in each of the sentences below;

Example:

Your letter (caused me to remember) that I should send you back your books.

A informed me
(B) reminded me
C announced to me
D suggested to me

(Note that the circle is made around the letter B because the words "reminded me" give the best meaning of the expression "caused me to remember".)

25. The parents were (feeling anxious) at the absence of their son.

A worried
B angry
C offended
D shocked

26. A nurse is a woman who (takes care of) children and sick people.

A likes
B dislikes
(C) looks for
D looks after

27. The King was very (well liked by the people).

A familiar
(B) popular
C humble
D attractive

28. This list does not (have in it) your name.

A consist
B tell
(C) include
D claim

29. I will (take as my son) the child, and I will be responsible for his living and his education.

A accept
B possess
(C) adopt
D invite

30. He (demanded to be given) the prize because he won the game.

A claimed
(B) took
C refused
D accepted

31. The (state of affairs) is becoming difficult.

(A) situation
B problem
C business
D system

III

Directions: Read the passage carefully; then choose the answers to the questions according to what the passage says.

Example:

"A good camera can take pictures at a high speed, but you have to pay too much for such a camera."

This paragraph tells you that a good camera is

(A) expensive
B cheap
C useless
D rare

(The only true answer is"expensive". Therefore the circle is made around A.)

Now read the following passage and answer the questions that follow it in the same way.

"When the bed was ready I lit a candle and went upstairs. I lay on the bed, leaving the candle burning, meaning to stay awake, but I was very tired and soon dropped off to sleep.

After a few hours I awoke suddenly; the door was opening of itself; I could see no one opening it. Then a cold wind blew out the candle, leaving me in black darkness. Through that inky darkness I heard the sound of a moan; then there was a dragging sound, as if someone was pulling a heavy sack across the floor along the passage outside.

At last I found the matches and lit the candle again. I looked outside in the passage; there was no one, but there on the floor near the doorway, I saw the dull red mark of a naked foot."

32. The moment I awoke I saw

A somebody opening the door.
B somebody putting out the candle.
(C) the door opening of itself.
D a sack at the door

33. I left the candle burning before I went to sleep because

(A) I meant to stay awake.
B I wanted to send the ghosts away.
C I was frightened.
D I forgot to put it out.

34. In the darkness I heard the sound of

A knocking at the door.
B tapping on the wall.
C heavy footsteps.
(D) a deep cry of pain.

35. When looking outside the room, I saw

A a man crossing the doorway.
(B) a footprint on the floor.
C somebody dragging a heavy sack.
D a stain of red blood.

36. By the expression "dropped off to sleep", the writer means

A waked up
B fell asleep
C fell down (from bed) while asleep
D fainted

37. By the expression "blew out the candle", the writer means

A put out the candle
B made the flame move sideways
C made the candle explode
D made the candle fall down

IV

Directions: Each of the following sentences is followed by four statements. Only one statement is true according to the information in the sentence. Make a circle around the letter in front of the true statement.

38. Tulips are flowers which are grown in Holland.

A Tulips have beautiful colours in Holland.
B Tulips are sold in Holland.
C Tulips are planted in Holland.
D The Dutch like Tulips.

39. Like birds, men have tried to fly with wings but they are too heavy.

A Men fly with wings.
B Men are like birds.
C Men are too heavy to fly.
D Birds are too light to fly.

40. The little drops of rain break up the white light of the sun into the rainbow colours.

A The colours of a rainbow are beautiful.
B Raindrops separate the sunlight into different colours.
C Raindrops are coloured.
D A rainbow consists of little drops of rain.

41. He left his horse in a hut and went on by a boat along the river.

A His horse was tired.
B He walked along the river.
C He slept in a boat.
D His horse was left in a hut.

V

Directions: Make a circle around the letter in front of the best English translation of each of the Arabic sentences:

42. دُعِيَ كثيرٌ من النّاس الى حفلةِ عشاءٍ في القصر

A Many people will be called to the dinner party in the Palace.
B Many people were invited to a dinner party in the Palace.
C Many people are requested to come to a dinner ceremony at the palace.
D Many people are going to the dinner ceremony in the Palace.

43. رجا السجينُ حارسَهُ أن لا يعاملَه بقسوة

A The prisoner told his guard to treat him not severely.
B The prisoner ordered his guard not to treat him hardly.
C The prisoner demanded that his guard should not treat him hardly.
(D) The prisoner asked his guard not to treat him severely.

Directions: Make a circle around the letter in front of the best Arabic translation of each of the English sentences:

44. We had to cancell the match because it was so foggy.

(A) كان علينا أن نُلغي المباراة لأنّ الضبابَ كان كثيفاً
B ستُؤجَّل المباراة اذا كان الضبابُ كثيفاً
C قد تجرى المباراة اذا تبيّن أنّ الضبابَ لن يكونَ كثيفاً
D لقد أُجِّلتْ المباراة لانّه تبيّنَ أنّ الطقسَ سوف يتحسّن

45. The man who entertained us in his house for two days was very generous.

(A) كان الرجلُ الذى احتفى بنا في بيتِهِ مدةَ يومين كريماً جداً
B كان الرجلُ متسامحاً حين استضافَنا في بيتِهِ مدةَ يومين
(C) الرجلَ الشجاعُ هو من يُكرِمُنا في بيتِهِ مدةَ يومين
D لقد استضافَنا الرجُلَ لمدةِ يومين لما كان له من منزلةٍ كبيرةٍ في بيته.

VI

Directions: Each group of lines of verse below is followed by four statements. Only one of the statements is true according to what meanings the verse gives. Make a circle around the letter in front of the true statement.

Example:

"But to go to school in a summer morn
O! it drives all joy away;"

The poet in these lines of verse says that going to school is

(A) joyful.
B useful.
C necessary.
(D) unpleasant.

(Statement D is true according to what meanings there are in the verse.)

46. "I sent thee late a rosy wreath,
Not so much honouring thee
As giving it a hope, that there
It could not withered be." (Ben Jonson)

A I sent you flowers so that you will love me more.
(B) I hoped that the flowers I sent you would not dry up.
C I thought that my flowers would please you.
D I sent you flowers that will give you hope.

47. " Give me the strength never to disown the poor or bend my knees before insolent might. "

(R. Tagore)

A Help me to overcome difficult tasks.
B Help me to be strong for 'might makes right'.
(C) Help me not to bow to an unkind man.
D Help the poor to bear their poverty.

48. " As one lamp lights another, nor grows less,
So nobleness enkindleth nobleness. "

(James R. Lowell)

(A) Goodness creates goodness.
B Plant good and you'll harvest evil.
C Lighting a lamp is a noble deed.
D When a lamp burns its light decreases.

49. " So when my Mistress shall be seen
In form and beauty of her mind,
By virtue first, then choice, a Queen,
Tell me, if she were not designed
Th' eclipse and glory of her kind? "

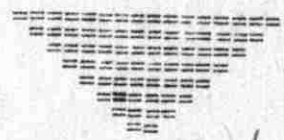
(Sir Henry Wotton)

(A) My Mistress shall be chosen a queen for her beauty.
B My Mistress wishes to be a queen of all women.
C The Queen has chosen my Mistress as her maid.
D My Mistress is wise and good.

50. " Alone she cuts and binds the grain,
And sings a melancholy strain;
O listen! for the vale profound
Is overflowing with the sound. "

(W. Wordsworth)

(A) She was singing a joyful song.
B She was crying in the deep valley.
C She was reaping by herself.
D She heard a deep sound coming from the valley.



0 : 1

$31 - \frac{6}{3} = 29$

$7\frac{1}{2}$

$36\frac{1}{2}$

A P P E N D I X IV

Experimental Objective Achievement Test

in

Arabic Language

في اللغة العربية

(مستوى امتحان الشهادة الاعدادية العامة)

للمدارس الاردنية

(21) 2023

الاسم : علي محمد علي الصور

الصف : الاول الثانوي

الزمن : خمسون دقيقة

تعليمات للاجابة على الاختبار

١ــ لا تقلب هذه الصفحة الا عندما يُعلن عن ذلك

٢ــ لا تبدأ الاجابة الا بعد ان تقرأ جيداً التعليمات المذكورة في هذه الصفحة وفي مستهلّ كلّ مجموعة من الاسئلة الواردة داخل الكراس

٣ــ حاول الانتباه لما تقولُه التعليمات والتقيّد بها عند الاجابة

٤ــ وُضّحت التعليمات ببعض الامثلة المُجاب عنها يمكن الاسترشادُ بها

٥ــ عدد الاسئلة في هذا الاختبار ستون ، وكذلك العلامة الكاملة على هذا الاختبار هي ستون ، اى أن لكل سؤال علامة واحدة .

٦ــ تُحسب علامتك على هذا الاختبار بعد خصم جزء من اجاباتك الخاطئة من مجموع اجاباتك الصحيحة ولذلك فليس من صالحك عندما لا تعرف الجواب الصحيح ان تخمّنه الا اذا استطعت ان تضيّق مجال اختيارك للجواب الصحيح .

٧ــ اذا بدا لك احد الاسئلة صعباً ويحتاج وقتاً للتفكير اتركه على ان تعودَ إليه فيما بعد إذا تبقّى لديك وقت حينذاك

٨ــ لا يتوقع ان يجيب اىّ منكم على جميع الاسئلة صحيحةً في الوقت المحدّد ، حاول فقد يحالفك الحظ .

لا تقلب هذه الصفحة الا عندما يُعلن عن ذلك .

(176) فيما يلي جُمل وُضع خطٌ تحت كلمةٍ في كلٍّ منها ، وقد كُتبت الحروف أ ، ب ، جـ ، د امام كل جملة . اصنع دائرة

حول الحرف أ اذا كانت الكلمة التي تحتها خط تُعرب "مفعولاً به"

" " " ب " " " " " " "تمييـــزاً "

" " " جـ " " " " " " "حـــالاً "

" " " د " " " " " " "مفعولاً مطلقاً "

مثـــال (أ) ب جـ د قرأت قصة ممتعة

(رسمت الدائرة حول أ للدلالة ان كلمة "قصة" تعرب مفعولا به)

١ـ ✗ (أ) ب جـ د سرّني رجوعك سالما

٢ـ ✗ أ ب (جـ) د اتّخذتُ الكتاب سميرا

٣ـ ✓ أ (ب) جـ د كم كتابا في مكتبتك ؟

٤ـ ✗ أ ب (د) جـ جاء الطالب المدعوّ خالدا

٥ـ ✓ أ ب جـ (د) اثّرت القصة في نفسي تاثيرا شديدا

٦ـ ✗ أ ب جـ (د) خالد اكبر منك سنا

٧ـ ✓ (أ) ب جـ د كناطحٍ صخرة يوما ليوهنها ...

٨ـ ✓ أ ب (جـ) د الى الله مرجعكم جميعا

القسم الثاني

فيما يلي كلمات كتب امام كل منها الحروف أ ب جـ د

ارسم دائرة حول الحرف أ اذا كانت الكلمة "اسم فاعل"

" " " ب " " " "اسم مفعول"

" " " جـ " " " "مصدراً"

" " " د " " " "صفة مشبهة"

٩ـ ✓ أ ب جـ (د) جيّد

١٠ـ ✓ أ ب (جـ) د موعِد

١١ـ ✓ أ ب (جـ) د تواضُع

١٢ـ ✓ أ (ب) جـ د مدعوّ

١٣ـ ✓ أ ب (جـ) د سطوة

١٤ـ ✓ (أ) ب جـ د عالِم

١٥ـ ✓ أ (ب) جـ د متّهم

١٦ـ ✗ أ ب جـ (د) رحيم

اقرأ الابيات التالية ، ثم اجب عن الاسئلة الخمسة التي تليها وذلك بان ترسم دائرة حول حول الحرف الذى يقع الى جانب الجواب الصحيح .

* ونشرب إن وردنا الماء صفواً ويشرب غيرنا كدراً وطينا
اذا ما الملك سام الناس خسفاً أبينا أن نقر الذل فينا
ملأنا البر حتى ضاق عنا وماء البحر نملأه سفينا
اذا بلغ الفطام لنا صبيٌّ تخر له الجبابر ساجدينا *

مثال : ما المعنى المقصود من العبارة " سام الناس خسفاً " ؟

أ رفع عنهم الظلم والحيف
(ب) أكرههم على ما فيه ذلهم
ج عدل بينهم
د حثهم على مقاومة من يظلمهم

(لقد رسمت الدائرة حول الحرف ب الذى يسبق الاجابة الصحيحة، وبنفس الطريقة اجب على الاسئلة التالية)

١٧ـ أي من عصور الادب العربي ينتمي اليه هذا الشعر ؟

أ عصر النهضة
ب العصر العباسي
ج العصر الاسلامي
(د) العصر الجاهلي

١٨ـ اى من اغراض الشعر يتمثل في الابيات المذكورة ؟

أ الهجاء
ب الغزل
(ج) الفخر
د الوصف

١٩ـ أيّ من العبارات التالية يتفق وغرض الشاعر في البيت الاول ؟

أ نحن اغنياء وغيرنا فقراء
(ب) نحن السادة فلنا من كل شيء افضله ولغيرنا ارذله
ج الماء عندنا انقى واصفى مما هو عند غيرنا
د نحن نحافظ على نظافتنا في المأكل والمشرب وغير ذلك اكثر من غيرنا

٢٠ـ في البيت الثاني يؤكّد الشاعر صفةً في بني قومه هي :
(أ) عزة النفس
ب الخنوع والذّلّة
ج الاخلاص لسيّد القوم وإن ظلم
د القسوة والتسلط على الضعيف

٢١ـ أي من العبارات التالية يتفق وغرض الشاعر من قوله " ملأنا البر حتى ضاق عنا "
أ قطعنا في مسيرنا مسافةً طويلة
ب خضنا مع خصومنا معارك كثيرة
(ج) ملأنا البرّ برجالنا فلم يتّسع لنا لكثرتهم
د قصدنا طلب الرزق فلم نُوفّق

القسم الرابع

اقرأ القطعة الادبية التالية ثم بيّن صحة أو خطأ الجمل التي تليها حسب ما فهمته من القطعة . لقد كتب الحرفان ص ، خ امام كلٍ من هذه الجمل وما عليك الا ان ترسم دائرة حول ص اذا كان معنى الجملة يتفق ومضمون القطعة . او ترسم الدائرة حول خ اذا كان معنى الجملة لا يتفق ومضمون القطعة .

" أغنياؤنا كثيرون ، ولكن معظمهم أشدّ بؤساً من الفقراء المعوزين لانهم لا يفقهون الثروة ولا يقدّرونها ، ولا يفهمون ما ينبغي ان توجد هذه الثروة من صلة بينهم وبين مواطنيهم . وهم اغنياء وكلّ حظّهم من ثروتهم ان ياكلوا كثيرا ويستمتعوا بلذات مادية لا تتجاوز الحسّ الى القلب او الى العقل . ثروتهم مقصورة على اجسامهم فان وصلت الى نفوسهم فهي لا تمسّ منها الا موضع الضعف والغرور ، تمس الفخر والتّيه ، تمسّ العُجْب والخُيَلاء ولكنها لا تمسّ الذكاء ، ولا تمس عاطفة الرحمة بالبائس ولا تمس عاطفة الاعانة على الخير "

مثال (ص) (خ) حاول الكاتب في هذه القطعة ان يطرق موضوعاً اجتماعياً باسلوبٍ أدبي

(لقد رسمت الدائرة في هذا المثال حول الحرف ص لانك لو تمعّنت في اسلوب وموضوع القطعة لاستنتجت صحة ما جاء في هذه الجملة)

٢٢ـ (ص) (خ) يمتدح الكاتب في هذه القطعة ما يقوم به اغنياء بلاده من بذل اموالهم في سُبل الخير

٢٣ـ (ص) (خ) ينتقد الكاتبُ في هذه القطعة الطريقةَ التي يُنفق فيها اغنياءُ بلادِه اموالَهم

٢٤ـ (ص) (خ) من رأي الكاتب أن هؤلاء الاغنياء يدركون جيّداً كيف تُستخدم الثروةُ في توطيد الصّلات الطيبةِ بينهم وبين مواطنيهم

٢٥ (ص) (خ) يرى الكاتب ان هؤلاء الاغنياء لا يُنفقون ثرواتهم على الملذات الحسّية وحسب ، وانما يقصدون ما فيه متعة للعقل او ما فيه عمل للخير

يسْعَوْن من وراء ذلك للحصول على شهرة زائفة او ليشبعوا شهوة الغرور في انفسهم .

القسم الخامس

العبارات التالية بعضُها صحيح وبعضها خطأ . فاذا كانت العبارة صحيحة ارسم دائرة حول ص ، وان كانت خطأ ارسم دائرة حول خ .

٢٧ـ ص خ تمتاز المعلقات ـ في الشعر الجاهلي ـ بان القصيدة منها تحوى غرضاً واحدا من اغراض الشعر تدور حوله كلُّ ابيات القصيدة .

٢٨ـ ص خ كثير من الآثار الادبية التي نُقلت الينا من النثر الجاهلي هي على هيئة حكم وامثال .

٢٩ـ ص خ بقيت مواضيع الشعر في العصر الاسلامي كما كانت عليه في العصر الجاهلي لم يستجد عليها جديد .

٣٠ـ ص خ يبرز في خطب الحجاج بن يوسف الثقفي التسامح واللين والاخذ بالمعروف

٣١ـ ص خ اشتهر في شعر الفرزدق الفخر والاعتزاز بالاصل والنسب

٣٢ـ ص خ كانت الطبيعة في الاندلس من اهم العوامل التي طبعت الشعر الاندلسي وميّزته عن شعر المشرق .

القسم السادس

اضْبط بالشكل الصحيح الحرفين الاخيرين من كل كلمة تحتها خط في الجمل الاتية .

اجعل علامات الشكل واضحة على النحو الذى يظهر في المثال التالي :

مثال لم يكتُبْ خالِدٌ رسالَةً الى صديقِهِ

٣٣ـ أيها الاغنياءُ لا تسرفوا في الخيلاءْ

٣٤، ٣٥ ان الله غيرُ تاركٍ عبيدَهُ يظلمون

٣٦ـ قولك الصدْقَ واجب

٣٧ـ ما جاء بك الساعةَ

٣٨ـ وجدت طائراً مقصوصا جناحَهُ

٣٩، ٤٠ ما عرفت احد اشرف نفساً او ارْقَ عاطفة من سعادٍ

القسم السابع

الكلمة المحذوفة في كل من الجمل الاتية هي احدى كلمتين يجوز في المناسبة المذكورة استعمال احداهما دون الاخرى . ارسم دائرة حول الحرف الذى يشير الى الكلمة التي يكون استعمالها في الفراغ صحيحا . (لا تكتب شيئا في الفراغ) .

نجح ـــــــــ في الامتحان

(ب) أخوك

(لقد رسمت الدائرة حول ب لان كلمة اخوك هي الكلمة التي يكون استعمالها صحيحا هنا)

٤١ـ أ ثماني
(ب) ثمانية
كان بالمرسى ـــــــــ سفن

٤٢ـ أ الأطول
(ب) أطول
عدنان ـــــــــ من اخيه

٤٣ـ (أ) جريح
ب جريحة
جاءتنا امرأة ـــــــــ

٤٤ـ أ سالمون
(ب) سالمين
رجع المسافرون الى بلادهم ـــــــــ

٤٥ـ (أ) بصره
ب بصيرته
اذا أصيب شخص بالعمى قيل أنه فقد ـــــــــ

٤٦ـ أ عَييتُ
(ب) أعْيَيْتُ
ـــــــــ في المشي

٤٧ـ (أ) غَلَقْتُ
ب أَغْلَقتُ
ـــــــــ الباب واقفلته

٤٨ أ رماك
(ب) أرماك
اذا ركبتَ الفرس ـــــــــ

٤٩ـ (أ) آزرتَهُ
ب وازرتَهُ
اذا أعنْتَ الشخصَ في أمرٍ قلتَ انك ـــــــــ

٥٠ـ أ واكلتَهُ
(ب) آكلتَهُ
اذا أكلتَ مع شخصٍ قلتَ أنك ـــــــــ

تعتبر الكتب التالية من الآثار الادبية المهمة عند العرب . اكتب اسم المؤلف لكل من هذه الكتب في الفراغ المعد الى جانبه .

	اسم الكتاب	المؤلف
٥١ـ	العقد الفريد	ابن عبد ربه الأندلسي
٥٢ـ	البخلاء	الجاحظ
٥٣ـ	الامالي	ابو علي القالي
٥٤ـ	المدينة الفاضلة	
٥٥ـ	كليلة ودمنة	عبد الله المقفع

القسم التاسع

ألّف من كل مجموعة من العبارات فيما يلي جملةً تامة المعنى دون ان تضيف اى كلمة من عندك ، ثم اكتب كلا من هذه الجمل في الفراغ المعد لها .

٥٦ـ (من أجْل) ، (الجزائريّون) ، (طويلاً) ، (استقلال) ، (ناضَل) ، بلادهم زمناً)

ناضل الجزائريون من أجل استقلال بلادهم زمناً طويلاً

٥٧ـ (لسانه لدى) ، (الدهشة) ، (النبأ) ، (عقدت) ، (سماع) .

عقدت الدهشة لسانه لدى سماع النبأ

٥٨ـ (الصالح يخدم) ، (على رفعة) ، (المواطن) ، (شأن وطنه) ، (امته ويعمل)

المواطن الصالح يخدم أمته ويعمل على رفعة شأن وطنه

٥٩ـ (المتحضرة في) ، (في العلوم) ، (تتميز الامم) ، (الحاضر بما) ، (العصر)
(تقدم) ، (احرزته من)

تتميز الأمم المتحضرة في العصر الحاضر بما أحرزته من تقدم في العلوم

٦٠ـ (السائحين) ، (الاصطياف التي) ، (إليها عدداً) ، (الاثرية ومناطق) ، (تجذب)
(كبيراً من) ، (يوجد في) ، (من الاماكن) ، (الاردن كثير)

يقصد السائحون من كثير من الأماكن في الأردن عدداً كبيراً من مناطق الاصطياف

A P P E N D I X V

Standardized Student Answers

on

The General Secondary Certificate Examination

(in Jordan)

Standardized Student Answer to a Question
on English Language (Sec. Exam.)

إجابة لطالب السؤالين الغائبين (مجهولين)

2 (A) Della and Jim were wise because they ~~a~~ sacrified by the best thing which every one have to get the suitable present To his husband, but last these presents appear such as the presents of the Indian Wise Men who offered their presents to Jusus when he was in the cred'et because they are useless ~~for The fob which brought Della~~

(4) because the watch and the tresses of ~~her~~ hair were lost and then their presents (the fob and the combs) were now useless.

4_ 1. Who is responible about this deed

(0) 2. He came at the 20th on May.

Zero 3. Students are fond ~~the~~ to be defetive stories.

4. The maid servant does't look a bout the baby.

اجابة السؤال الثالث في التاريخ لطالب (مجهول)

السؤال: اشرح باختصار الوسائل التي لجأ اليها الانتداب البريطاني في تهويد فلسطين

Standardized Student Answer to a
Question on History (Sec. Exam.)

اسس التدقيق

١- مجرد ذكر كل من النقاط التالية ينال كل منها ٥ر علامة:
١- الهجرة ٢- بيع الاراضي ٣- الحالة الاقتصادية (تشجيع الاقتصاد اليهودي وخنق الاقتصاد العربي) ٤- الامتيازات لليهود ٥- التسليح

٢- العلامة الكاملة لكل نقطة تذكر وتشرح (٤) علامات

٣- اذا ذكرت نقاط اخرى (كالادارة مثلاً) تعطى علامة تضاف الى العلامات

٤- العلامة الكاملة للسؤال هي (٢٠)

الاجابة

س٣ الوسائل التي لجأت اليها بريطانيا في تهويد فلسطين هي

① الهجرة
② الاراضي
③ تسليح اليهود
④ ~~الحركات~~ الامتيازات الاجنبية

① الهجرة
عندما فكرت بريطانيا بانشاء وطن قومي لليهود في فلسطين اخذت تعمل على تسهيل الهجرة اليهودية الى بعض مناطق فلسطين تدريجيا حتى انه في النهاية اخذت تخبر الحكومة بأنها سوف لا تدع اي يهودي بالدخول الى فلسطين من غير امر او رضاء العرب عن ذلك ولكنها كانت تسمح لهم عن طريق التهريب باعداد هائلة حتى اصبحت مزدحمة ازدحاما شديدا بهم فما بقي الا ان اطردوا العرب من مساكنهم وحلوا بها وقد بلغ عددهم فوق ما يتصور الانسان في النهاية وكانت بسبب ذلك كما تقول بريطانيا اللعينة انهم مساكين ضعفاء لا يجدون لهم بلادًا يأوون اليه اكثر من فلسطين

② الاراضي
سمحت بريطانيا للعرب ببيع اراضيهم الى اليهود لكي يشتغلوا بها ... متنوعات حتى اصبح الكثير من العرب فاقدا

من بريطانيا بعدم بيع الاراضي لليهود فردت على ذلك بانها
سوف تمنع بيع الاراضي بعد اكتفاء اليهود مما يحتاجونه الخ.

(٣) تسليح اليهود

فقد قامت ~~في~~ بريطانيا بالاعمال التالية من جهة التسليح :-

١) تدريب اليهود تدريبا عسكريا
٢) تعريف اليهود بما يسمونه الاسلحة الحديثة وكيفية استعمالها
٣) انشاء مستعمرات بوليس سري
٤) امدادهم بعدد كبير من الاسلحة فقد بلغت اول الامدادات من الاسلحة كما يلي بندقية ١٧٢٧ رشاش ١٧ فرد ٣٠ دبابة ١١
٥) تهريب الاسلحة اليهم من موانيها التي كان يحرسها اليهود انفسهم باعداد هائلة حتى اصبحوا هم يفوقون القوة العربية التي كانت فقيرة الحال من جهة الاسلحة.

(٤) ~~المشروعات~~ الامتيازات

أ) مشروع كهرباء روتنبرج اليهودي فقد اعطى هربرت صموئيل لليهودي امتياز المشروع كهرباء وتتضمن ما يلي : ١) مدة الامتياز ٧٥ عاما ٢) لصاحب المشروع ان ~~[illegible]~~ يستثمر نهر الاردن وروافده ٣) له الحق في انارة اي بلد في فلسطين ما عدا القدس ٤) ~~[illegible]~~ له الحق في استثمار مياه نهر العوجا ٥) للحكومة الذي يستقل بالمشروع حق تغيير مجرى نهر اليرموك

ب) مشروع البحر الميت :-

ويتضمن ما يلي :-

١) انشاء شركة اسمها " شركة البوتاس الاردنية "
٢) مدة الامتياز ٧٥ عاما
٣) استثمار جميع ما يحويه البحر الميت من املاح ومعادن وغير ذلك

ج) مشروع تجفيف الحولة :-

فقد اعطى هذا المشروع اولا لشخص لبناني لاستثمار عدد كبير من آلاف الدونمات من الاراضي عند بحيرة الحولة ولكنه عجز عن ~~[illegible]~~ اقامته فدفع له شخص يهودي مبلغ ٢٣٠ الف دينار على اساس انه يتخلى عن المشروع فقبل واشترى هذا الشخص اليهودي ~~[illegible]~~ في بداية الانشاء
[illegible] بريطانيا انه [illegible] الوطن

Standardized Student Answer to a Question on Philosophy (Sec. Exam.)

٩ تسع علامات

اجابة السؤال الأول في الفلسفة للطالب (مجهول)

القيم المطلقة وهي التي تكون صالحة لكل زمان ومكان وتكون صالحة لجميع الاشخاص وعلى جميع المزاجات وهي ~~قيم الجمال~~ كما حددها افلاطون قيم الجمال ويدرسها علم الجمال وقيم الحق ويدرسه علم المنطق وقيم الخير ويدرسه علم الاخلاق

والفرق بينها وبين القيم النسبية انها شائعة ~~تحكم~~ احكام واوصاف تعطى على الاشياء كل حسب مزاجه فهي شخصية تتأثر بالاحساسات الفردية ~~[illegible]~~ اي انه اي فرد يستطيع ان يعطي قيمة مثلا ثم موضوعي - اي شيء ربما تكون قيمتها بالنسبة لشخص وتكون لآخر سيئا. وكذلك ظرفية او تتأثر بالاماكن فمن الممكن ان يستعمل جديد لفظ يعطي قيما احسن واجود مما يستعمل ~~[illegible]~~ من الله قديم وكذلك يتأثر بالزمان اذ ان القيمة التي تكون للفرد في وقت زخارفها تختلف ~~[illegible]~~ عما كانت عليه قبل وتكون الآن اليها تختلف.

وللقيم النسبية ثلاث اطراف ~~[illegible]~~ الاول متعلق بالشخص المعطي للقيمة والثاني بالشيء الذي يراد اعطائه القيمة والثالث ~~[illegible]~~ للموضوع الشيء.

مثل ان الاعجاب بتمثال لمايكل ~~[illegible]~~ انجلو واعطاءه القيمة الحسنة يرجع الى الاعجاب الشخصي نفسه بما رأى من فن واتقان في هذا التمثال وكذلك يرجع الى التمثال نفسه بان يكون به الجمال ليكون بين جوانبه من الدقة في الفن كما يكون الفنان الذي نحت التمثال. والقيم تتغير وتتبدل حسب الظروف والامكنة والقيم النسبية انواع كثيرة منها:

١- قيم اخلاقية مثل مؤدب رذيل
٢- قيم دينية مثل حلال حرام
٣- قيم جمالية رائع قبيح
٤- قيم تتعلق بالصحة عليل سقيم
٥- قيم قانونية مشروع وغير مشروع

السؤال الثاني من الورقة الثانية لفحص الرياضيات

أ) إذا كان م ، ن جذري المعادلة أ س^2 + ب س + د = صفر ، أوجد المعادلة التي جذراها $\frac{1+م}{1-م}$ ، $\frac{1+ن}{1-ن}$. (١٢ علامة)

ب) برهن على أن جذري المعادلة س^2 - ٢ك س + ٢ك - ٢ = صفر تخيليان لجميع قيم ك الواقعة بين ١ ، ٢ . (١٨ علامة)

جواب الطالب (مجهول)

مجموع الجذرين للمعادلة أ س^2 + ب س + د = صفر

$م + ن = -\frac{ب}{أ}$

$م ن = \frac{د}{أ}$

مجموع الجذرين للمعادلة المطلوبة

$$\frac{1+ن}{1-ن} + \frac{1+م}{1-م} = \frac{(1-ن)(1+م) + (1+ن)(1-م)}{(1-ن)(1-م)}$$

$$= \frac{1+م-ن-م ن + 1+ن-م-م ن}{1-ن-م+م ن} = \frac{2-2م ن}{1-(م+ن)+م ن}$$

$$= \frac{2-2\frac{د}{أ}}{1+\frac{ب}{أ}+\frac{د}{أ}} = \frac{2-\frac{2د}{أ}}{\frac{أ+ب+د}{أ}} = \frac{2أ-2د}{أ+ب+د}$$

مجموع الجذرين

حاصل ضرب الجذرين

$$\frac{1+م}{1-م} \times \frac{1+ن}{1-ن} = \frac{1+م+ن+م ن}{1-ن-م+م ن} = \frac{1+(م+ن)+م ن}{1-(م+ن)+م ن}$$

$$= \frac{1+\frac{ب}{أ}+\frac{د}{أ}}{1-\frac{ب}{أ}+\frac{د}{أ}} = \frac{\frac{أ+ب+د}{أ}}{\frac{أ-ب+د}{أ}} = \frac{أ+ب+د}{أ-ب+د}$$

حاصل ضرب الجذرين

المعادلة س^2 - $\left(\frac{2أ-2د}{أ+ب+د}\right)$ س + $\frac{أ-ب+د}{أ+ب+د}$ = صفر ×

= (أ + ب + د) س^2 - (٢أ - ٢د) س + (أ - ب + د) = صفر

وهو المطلوب

ب) س^2 - ٢ك س + ٢ك - ٢ = صفر

$ب^2 - ٤أ د = (-٢ك)^2 - ٤ \times ١ \times (٢ك - ٢)$

٤/ ٤ب² − ١٢ب + ٨ = ٠

ب² − ٣ب + ٢ = صفر

(ب − ٢)(ب − ١) ∴ ب = ٢ ، ١

جميع قيم ب الحقيقية الواقعة بين ١ ، ٢ لأنه لا تأخذ إلا القيم ١ ، ٢

فيما عدا ذلك خارجها تخيلية

الاجابة الصحيحة وتوزيع العلامة على اجزاء الاجابة

أ) م + ن = − ب/أ ، م ن = حـ/أ

(١+م)/(١−م) + (١+ن)/(١−ن) = (١+م−ن−م ن + ١+ن−م−م ن) / ((١−م)(١−ن))

= (٢ − ٢م ن) / (١ − ن − م + م ن)

= (٢ − ٢م ن) / (١ − (م+ن) + م ن) = (٢ − ٢حـ/أ) / (١ + ب/أ + حـ/أ) = (٢أ − ٢حـ) / (أ + ب + حـ) — ٥ علامات

(١+م)/(١−م) × (١+ن)/(١−ن) = (١+ن+م+م ن) / (١−ن−م+م ن) = (١+(م+ن)+م ن) / (١−(م+ن)+م ن)

= (١ − ب/أ + حـ/أ) / (١ + ب/أ + حـ/أ) = (أ − ب + حـ) / (أ + ب + حـ) — ٥ علامات

المعادلة س² − ((٢أ − ٢حـ)/(أ+ب+حـ)) س + (أ−ب+حـ)/(أ+ب+حـ) = صفر — ٢ (علامتان)

(أ+ب+حـ)س² − (٢أ − ٢حـ)س + (أ−ب+حـ) = صفر — ١ (علامة)

١٣ علامة

ب) س² − ٢ب س + (٣ب − ٢) = صفر

المميز ب² − ٤أحـ = ٤ب² − ٤(٣ب − ٢) — ٢ (علامتان)

المعادلة المناظرة للمميز ٤ب² − ٤(٣ب − ٢) = صفر

(ب² − ٣ب + ٢) = صفر

(ب − ٢)(ب − ١) = صفر — ٤ علامات

∴ ب = ٢ ، ب = ١

هذا المقدار موجب دائماً إلا إذا وقعت قيمة ب بين ١ ، ٢ — ٦ علامات

[illegible]

Standardized Student Answer to a Question
on Physics (Sec. Exam.)

الفيزياء

اجابة السؤال الثالث للطالب (مجهول)

أ : $ح = هـ ظا ح = \frac{ع ٢}{ف ٣}$

في حالة المغناطيس القصير في الوضع الأول لجاوس

ح = شدة المجال المغناطيسي ، هـ : المركبة الافقية للمجال الارضي

ح : زاوية الانحراف ، ع : عزم المغناطيس ، ف : البعد عن مركز الارض

نأخذ المغناطيس ونضعه في مستوى الزوال المغناطيسي بحيث يكون ذا اتجاه شمال جنوب

نضع المغناطيس - أ - على احد الذراعين ونجد الانحراف الذي يطرأ على الابرة ونقرأه بدقة . ولتكن الزاوية ح، ثم نضع نفس المغناطيس على نفس البعد من مركز الابرة على الذراع الثاني ونجد الانحراف نكرر العمل عدة مرات ونقيس المسافة ولتكن ف، .

ثم نأخذ المغناطيس - ب - على احد الذراعين ونجد الانحراف الناتج من تأثير قطبه ثم نعكس الوضع ونكرر العمل عدة مرات ونأخذ معدل الانحراف ونقيس المسافة ولتكن الزاوية ح ، والمسافة ف ،

$هـ ظا ح_١ = \frac{ع_١ ٢}{ف_١^٣}$ ، $هـ ظا ح_٢ = \frac{ع_٢ ٢}{ف_٢^٣}$

$\therefore \frac{هـ ظا ح_١}{هـ ظا ح_٢} = \frac{ع_١ ٢}{ف_١^٣} \times \frac{ف_٢^٣}{ع_٢ ٢}$

وبمعرفة المسافة وظل الزاوية يمكن معرفة نسبة العزمين ، فاذا وضعنا المغناطيسين على نفس البعد في كل حالة فان

$\frac{ظا ح_١}{ظا ح_٢} = \frac{ع_١}{ع_٢}$

ب $ع_١ = ١٠٨$
$ع_٢ = ١٩٢$

نفترض وجود قطب شمالي مفرد في نقطة التعادل

∴ ١٠٨ = ٢ش١ ∴ ش١ = ٥٤ وحدة

١٩٢ = ٢ش٢

ش٢ = ٩٦ وحدة

∴ حـ من المغناطيس الأول = $\frac{١ \times ٥٤}{٩٠ \times ٣٠}$ = $\frac{٢}{٢}$ جاوس

حـ = = الثاني = $\frac{٩٦}{٤٠ \times ٤٠}$ = ٠٫٠٦ جاوس

وهذان المجالان متعامدان

∴ ح = $\sqrt{\frac{٤}{٩٠٠} + \frac{٣٦}{١٠٠٠٠}}$ = $\sqrt{\frac{٣٢٤ + ٤٠٠}{٩٠٠٠٠}}$

= $\frac{١}{٣٠٠}\sqrt{٧٢٤}$ = $\frac{٢٦٫٩}{٣٠٠}$ = $\frac{٢٧ \times ١}{٣٠٠}$ = ٠٫٠٨٧

∴ محصلة المجالين = ٠٫٠٨٧ جاوس

أسس الإجابة الصحيحة

أ - التجربة (طريقتان أي منها تأخذ ٣ علامات)

مقارنة العزمين بالاعتماد على القوانين التي يحفظها ٣ علامات

ب - المسألة ٨ علامات

علامتان للطالب الذي يجد شدة المجال الناشئ عن الأول وحده وهو ٠٫٠٠٨ اورستد

وعلامتان لمن يجد شدة المجال الناشئ عن الثاني وحده وهو ٠٫٠٠٦

والأربع علامات الباقية أي:

٣ علامات اذا توصل للقيمة $\sqrt{(٠٫٠٠٨)^2 + (٠٫٠٠٦)^2}$

علامة واحدة اذا استخرج الجواب النهائي.

A P P E N D I X VI

Standardized Student Answers

on

The General Preparatory Certificate Examination

(in Jordan)

Standardized Student Answer to a Question
on Arabic Language . (Prep. Exam.)

حضرة الزميل الكريم
فيما يلي اجابة طالب "مجهول" على السؤال (ب) من الادب والنصوص من ورقة اللغة العربية ، والاجابة المبينة هنا هي كما ظهرت في كراس الاجابات للطالب . الرجاء تقدير علامة لهذه الاجابة حسب الاسس التي تقوم عليها وهذه مذكورة مع نص السؤال في اسفل هذه الصفحة .
(مراعاة لامانة البحث العلمي ولاعتبارات اخرى الرجاء كتابة العلامة التي قدرتها)

الاجابة العلامة [١٢]

ب- قائل هذه الابيات الفرزدق
قالها في مناسبة : انه في ليلة كان ساهرا وقد جلس يتعشى فاذا بذئب يقترب على طعامه فقال للذئب اقدم فانه ان عاهدتني ان لا تخونني فنكون اصحابا فتأكل معي من طعامي وتسكن في بيتي حتى لا تخونني ولقد كان يتخيل تخيل الفرزدق
العواطف ولا الخيال
اذ اتى ذئب ، وقد كان جوعانا ولم يكن صاحبا فناديته ان يأتي لكي يتعشى انا واياه من هذا الطعام فلما قرب مني وقد اتى الي على ضوء النار وانا احمل الى عشى لتعشى قلت حين قرب تعالى الى هنا واقعد عندي ونكون في الزاد مشتركين ونأكل سواسية فاصبحت اقطع الزاد وآكل وهو يأكل على ضوء ضعيف ودخان كثير وهو يضحك واذا كنت متمكن بالعدة الكافية ان دنى مني لأقتلني دائما ولا يقع بيننا اي شيء ونكون مثل الاخوان .

نص السؤال :

ب - وأطلس عسال وما كان صاحباً : دعوت بناري موهنا فأتاني
فلما دنا قلت ادن دونك انني : وإياك في زادي لمشتركان
فبت أقد الزاد بيني وبينه : على ضوء نار مرة ودخان
فقلت له لما تكشر ضاحكا : وقائم سيفي من يدي بمكان
تعش ، فان عاهدتني لا تخونني : نكن مثل من يا ذئب يصطحبان

(١) من قائل هذه الابيات وفي اية مناسبة قيلت ؟
(٢) ضع عنوانا مناسبا لهذه القطعة واشرحها باسلوبك الخاص .

الاسس التي تقوم عليها في تقدير العلامة :

نقاط الاجابة : القائل (الفرزدق) ٣ علامات
المناسبة (الفرسلا) ٣ علامات
وضع عنوان مناسب ٣ علامات (يترك للمصحح تقدير العنوان المناسب)
شرح الابيات ٦ علامات
يراعى في شرح الابيات استيفاء معاني الابيات ، الاسلوب وصحة التعبير

Standardized Student Answer to a Question
on English Language (Prep. Exam.)

حضرة الزميل الكريم

فيما يلي اجابة طالب "مجهول" على السؤال الثالث (الترجمة) من ورقة اللغة الانجليزية. والاجابة المبينة هنا هي كما ظهرت في كراس الاجابات للطالب. الرجاء تقدير علامة لهذه الاجابة حسب الاسس التي تقوم عليها وهذه مذكورة مع نص السؤال في اسفل الصفحة.

(مراعاة لامانة البحث العلمي ولاعتبارات اخرى الرجاء كتابة العلامة التي تقدرها)

الاجابة	العلامة	
		٧

(A) ١- اغلبية البيوت تبنى من الحجر.

(C) ٢ - الدراجة التي اهداني اياها عمي هي هدية ميلادي

(d) ٤ - هل تعلمت كيف تكتب رسالة جيدة؟

(E) ٥ - إذا لم يأتي في الحال، وإلا سيجد الباص ذاهبا

(b) ٣ - من أين تعلم على الآلة الكاتبة؟

نص السؤال:

3. Translate the following sentences into Arabic:
(a) Most houses are built of stone (b) Where did he get the typewriter from? (c) The bicycle which my uncle gave me was a birthday present (d) Have you learned how to write a good letter? (e) Unless he comes at once, he will miss the bus.

الاسس التي تقوم عليها في تقدير العلامة

علامة السؤال = ١٠

توزع العلامة على خمس جمل لكل منها علامتان اذا كانت الجملة صحيحة. اما اذا ظهرت اخطاء في اي منها فتحسب العلامة حسب تقدير المصحح.

Standardized Student Answer to a Question
on Social Studies (Prep. Exam.)

حضرة الزميل الكريم

فيما يلي اجابة طالب "مجهول" على السؤال الاول في الرزمة الوطنية من ورقة الاجتماعيات ، والاجابة المبينة هنا هي كما ظهرت في كراس الاجابات للطالب . الرجاء تقدير علامة لهذه الاجابة حسب الاسس التي اتفق عليها وهذه مذكورة مع نص السؤال في اسفل هذه الصفحة .

(مراعاة لامانة البحث العلمي ولاعتبارات اخرى الرجاء كتابة العلامة التي قدرتها)

الاجابة

العلامة [٨,٥]

س١ ١) تتكون مجلس الأمة من البرلمان وتتكون البرلمان من مجلسين هما مجلس الاعيان ومجلس النواب .

٢) أ) عدد اعضاء مجلس النواب خمسون عضواً
ب) " " " الاعيان خمسة وعشرين عضواً .

٣) يعرض احد النواب مشروعاً فيدرسه فاذا وجده صالحاً عرضه على مجلس الاعيان فيدرس هذا المجلس ايضاً فان وافق عليه يقدم الى جلالة الملك ليصدق عليه وعندها يصبح قانوناً نافذاً .

نص السؤال : ممّ يتكون مجلس الامة ؟ اذكر عدد اعضاء كل من مجلسي الاعيان والنواب وبيّن كيفية عرض مشروع القانون والخطوات التي يتخذها حتى يصبح قانوناً نافذاً .

الاسس التي اتفق عليها في تقدير العلامة :

علامة الاجابة الكاملة ١٦,٥

الاجابة قسمين : الاول ١) اجابة ممّ يتكون مجلس الامة : النواب ، الاعيان
٢) عدد اعضاء كل من المجلسين : ٥٠ النواب ،
٢٥ الاعيان (اذا ذكر ان عدد اعضاء مجلس النواب ٦٠ تعتبر هذه الاجابة صحيحة)

علامة القسم الاول : (٨)

الثاني : عرض مشروع قرار ٨,٥ علامة
ويشمل ١) البرلمان له الحق في رفضه او تعديله او اقراره
٢) الاعيان : نفس الحق لكنه بعد اقرار النواب
٣) جلالة الملك : اقراره او رفضه .
٤) السلطة التنفيذية : لتنفيذه
اذا ذكرت هذه المراحل بدون تسلسل فيعتبر للاجابة علامتان (من ٨,٥) فقط .

حضرة الزميل الكريم

فيما يلي اجابة طالب (مجهول) على السؤال التاسع في مادة العلوم كما ظهرت في كراس الاجابات للطالب. الرجاء تقدير علامة لهذه الاجابة حسب الاسس التي اتفق عليها وهذه الاسس مع نص السؤال مذكورة في اسفل الصفحة.

مراعاة لامانة البحث العلمي وللاعتبارات اخرى الرجاء كتابة العلامة التي قدرتها

العلامة [ ]

الاجابة

٩ أ: أكمل المعادلتين

٢,٥ $2NaNO_3 + H_2SO_4 \xrightarrow{\Delta} Na_2SO_4 + 2HNO_3$

٢ $2KClO_3 \xrightarrow{\Delta} 2KCl + 3O_2$

ب

٢٨٤ اكسجين / $2KClO_3 \longrightarrow 2KCl + 3O_2$

(٣٩,١ + ٣٥,٥) ١٦

٧٤,٦ ٣٢

؟ ٢٨٤

١٩٧,٢ ⟶ ١٦ ٧٤,٦

؟ ٢٨٤

$= \frac{284 \times 197.2}{16}$ = ٤٧٢٢,٨

الجواب (٤٧٠٢٥/٥)

٣٩,١ / ٧٤,٦ / ٣٥,٥ / ٧١ / ١٤٨,٢ / ١٩٧,٢ / ٤٧٢٢,٨

نص السؤال ٩ : أ- أكمل المعادلتين التاليتين:

١ $NaNO_3 + H_2SO_4 \xrightarrow{\Delta}$

٢ $KClO_3 \xrightarrow{\Delta}$

ب- مستعينا بالمعادلة الثانية احسب الكمية اللازمة من كلورات البوتاسيوم للحصول على ٢٨٤ غم من الاكسجين اذا عرفت ان الوزن الذري للاكسجين ١٦ وللبوتاسيوم ٣٩,١ وللكلور ٣٥,٥

الاسس التي اتفق عليها في تقدير العلامة:

① علامة الاجابة الكاملة = ١٠

② ٤ علامات اذا اكملت المعادلة الاولى بحيث: النواتج صحيحة والتوزين صحيحاً

$2NaNO_3 + H_2SO_4 \xrightarrow{\Delta} Na_2SO_4 + 2HNO_3$

اما اذا كان توزين المعادلة خاطئاً بينما النواتج صحيحة فتعطى لهذا الجزء ثلاث علامات فقط

واذا كان احد النواتج مذكوراً فقط فيعطى لهذا الجزء علامة واحدة فقط

③ ٣ علامات اذا اكملت المعادلة الثانية بحيث النواتج صحيحة والتوزين صحيحاً، وعلامتان فقط اذا كان التوزين خاطئاً والنواتج صحيحة. $(2KClO_3 \xrightarrow{\Delta} 2KCl + 3O_2)$

④ ٣ علامات اذا حسبت كمية كلورات البوتاسيوم اللازمة بحيث كان الجواب والحل صحيحين

[illegible]

Standardized Student Answer to a Question
on Sciences (Prep. Exam.)

حضرة الزميل الكريم

فيما يلي اجابة طالب (مجهول) على السؤال التاسع في مادة العلوم كما ظهرت في كراس الاجابات للطالب . الرجاء تقدير علامة لهذه الاجابة حسب الاسس التي اتفق عليها وهذه الاسس مع نص السؤال مذكورة في اسفل الصفحة .
مراعاة لامانة البحث العلمي ولاعتبارات اخرى الرجاء كتابة العلامة التي قدرتها

العلامة: ٧

الاجابة

٩ ا) الحل الكمال ليته

$2NaNO_3 + H_2SO_4 \xrightarrow{\theta} Na_2SO_4 + 2HNO_3$

$2KClO_3 \xrightarrow{\theta} 2KCl + 3O_2$

ب) ٢٨٤ الكسجين / $2K|Cl|O_3 \longrightarrow 2KCl + 3O_2$

~~(٢٩١ + ٢٥٫٥)~~ ~~٧٤٫٦~~ ~~١٦~~ ~~٢٨٤~~

١٩٧٫٢ ⟵ ١٦ ~~٧٤٫٦~~
؟ ٢٨٤

$= \frac{٢٨٤ \times ١٩٧٫٢}{١٦}$ = ٤٧٢٢٫٨

الجواب
(٢٧٠٫٢ ٢/٥)

٧٤٫٦ ، ٢٩٫١ ، ٢٥٫٥ ، ٧١ ، ١٩٧٫٢٢ ، ١٩٧٫٢

نص السؤال ٩ : اكمل المعادلتين التاليتين:

١ $NaNO_3 + H_2SO_4 \xrightarrow{\theta}$

٢ $KClO_3 \xrightarrow{\theta}$

مستعينا بالمعادلة الثانية احسب الكمية اللازمة من كلورات البوتاسيوم للحصول على ٢٨٤ غم من الاكسجين اذا عرفت ان الوزن الذري للاكسجين ١٦ والبوتاسيوم ٢٩٫١ والكلور ٢٥٫٥ .

الاسس التي اتفق عليها في تقدير العلامة : ① علامة الاجابة الكاملة = ١٠

② ٤ علامات اذا اكملت المعادلة الاولى بحيث : النواتج صحيحة والتوزين صحيحاً
$2NaNO_3 + H_2SO_4 \xrightarrow{\theta} Na_2SO_4 + 2HNO_3$
اما اذا كان توزين المعادلة خاطئا بينما النواتج صحيحة فتعطى لهذا الجزء ثلاث علامات فقط
واذا كان احد النواتج مذكوراً فقط فيعطى هذا الجزء علامة واحدة فقط

③ ٣ علامات اذا اكملت المعادلة الثانية بحيث النواتج صحيحة والتوزين صحيحاً ، وعلامتان فقط اذا كان التوزين خاطئاً والنواتج صحيحة . $(2KClO_3 \xrightarrow{\theta} 2KCl + 3O_2)$

④ ٣ علامات اذا حسبت كمية كلورات البوتاسيوم اللازمة بحيث كان الجواب والحل صحيحين
وعلامة فقط اذا كان ...

Standardized Student Answer to a Question on Mathematics

حضرة الزميل الكريم

فيما يلي اجابة طالب "مجهول" على السؤال الثاني من ورقة الرياضيات ، والاجابة المبينة هنا هي كما ظهرت في كراس الاجابات للطالب . الرجاء تقدير علامة لهذه الاجابة حسب الاسس التي تقوم عليها وهذه مذكورة مع نص السؤال في اسفل الصفحة ، " مراعاة لامانة البحث العلمي ولاعتبارات اخرى الرجاء وضع العلامة التي قدرتها "

٩	العلامة

الاجابة

الربح = المبلغ × السعر × الزمن

الربح = ٤٠٠٠ × ٢/١٠٠ × ٣ = ٢٤٠ دينارا الربح البسيط

الربح المركب = المبلغ × السعر

= = = ٤٠٠٠ × ٢/١٠٠ = ١٢٠ دينار

٤٠٠٠ + ١٢٠ = ٤١٢٠ دينار يصبح المبلغ

٤١٢٠ × ٢/١٠٠ = ١٢٢,٦ دينار

١٢٠ + ١٢٢,٦ = ١٤٢,٦ دينار ربح في سنتين

٢٤٠ - ١٤٢,٦ = ٩٦,٤ دينار الفرق فيكون افضل له لو وضع المبلغ في البنك ليربح ربحا بسيطا بسعر ٢٪ ولمدة ثلاثة سنين .

نص السؤال : ايهما افضل لرجل ان يضع مبلغ ٤٠٠٠ دينار في بنك ليربح ربحا بسيطا بسعر ٢٪ لمدة ثلاث سنوات ، او ان يضعه في بنك آخر ليربح ربحا مركبا بسعر ٢٪ لمدة سنتين ؟

الاسس التي تقوم عليها لتقدير العلامة

٣ علامات : ٤٠٠٠ × ٢ × ٣ / ١٠٠ = ٢٤٠ دينار الربح البسيط

٣ علامات : جملة الدينار بالفائدة المركبة لمدة سنتين = (١,٠٢)^٣ = ١,٠٦٠٩

٣ علامات : جملة المبلغ كله = = = = = ١,٠٦٠٩ × ٤٠٠٠ = ٤٢٤٢,٦

٣ علامات : الفائدة المركبة للمبلغ = ٤٢٤٢,٦ - ٤٠٠٠ = ٢٤٢,٦

علامتان ∴ افضل له الحالة الثانية اي الفائدة المركبة

---

١٤ علامة

اذا ظهرت الاجابة بطريقة اخرى صحيحة تؤدي للجواب الصحيح ... اعط بعض خطوات او اجزاء هذه

BIBLIOGRAPHY

Baron, D., and Bernard, H. W. Evaluation Techniques for Classroom Teachers. McGraw Hill Co. Inc. New York, 1958.

Downie, N. M., and Heath, R. W. Basic Statistical Methods. Harper & Brothers, Publishers, New York, 1959.

Gulliksen, H. Theory of Mental Tests. New York, John Wiley & Sons, Inc., 1950.

Monroe, W. S. (et. al.) Educational Tests and Measurements. Houghton Mifflin Company, New York, 1924.

Monroe, W. S. (ed). Encyclopedia of Educational Research. The Macmillan Company, New York, 1952.

Nunnally, J. C. Tests and Measurement, Assessment and Prediction. McGrow Hill Book Co. Inc. New York, 1959.

Ross, C. C. Measurement in Today's Schools. New York, Prentice-Hall, Inc. 2nd edition, 1953.

Thorndike, R. L., and Hagen, E. Measurement and Evaluation in Psychology and Education. New York, John Wiley & Sons, Inc., 2nd edition, 1961.

Wrightstone, J. W. (et. al.) Evaluation in Modern Education. American Book Company, New York, 1956.

Documents

Jordan, Ministry of Education. The Law of Education No. 20 of 1955. (in Arabic).

Jordan, Ministry of Education. The By-Laws of General Education, No. 56 of 1960. (in Arabic).

Jordan, Ministry of Education. The System of the General Preparatory Certificate Examination for 1959. (in Arabic).

Jordan, Ministry of Education. The System of the General Secondary Certificate Examination for 1961. (in Arabic).