The extent of the achievement tests quality in lightened of the total quality standards (survey study at Afif College of education)

Dr. Amani Abdel Rahman Makkawi Abdel Rahman
Assistant Professor of Curriculum and Teaching Methods
Faculty of Education, Sudan University of Science and Technology

Dr. Nadia Elzein Ahmed Dawina
Assistant Professor of Curriculum and Methods of Teaching English
Language Faculty of Education Afif - Shaqra University - Saudi Arabia

Dr. Shahinaz Abdelrahman Osman Bashir
Shagra university, College of Science and humanitarian studies

Abstract:
The aim of the study is to identify the extent of the achievement tests quality by investigating diversity of the test specifications in the Faculty of Education in Afif, the awareness of staff about the standards of good and quality testing and the level of commitment to those standards by the Faculty members while preparing their test papers. The study used the descriptive analytical method. The research tool used to collect data was a questionnaire to survey staff members opinions. A random sample of 46 staff members were chosen randomly to respond to the questionnaire. Furthermore, the study used the SPSS program to analyze the data using several statistical tests such as Pearson coefficient correlation and Alpha Cronbach to find the validity and reliability of the questionnaire. Mean and standard deviation was also calculated. Analyzing data, It was found that the tests contents prepared by the staff members are varied properly, the awareness of faculty staff members at the Faculty of Education Afif in terms of the standards of good and quality test standards was medium, the level of commitment to the quality testing standards was high and their level of commitment to quality specifications in their test papers was high.

Key Word: Total quality standards - Achievement tests- Good and quality testing
**Introduction:**
Tests are the most important measurement tools that can be used in evaluating students in different universities. In order to evaluate the level of student’s achievement in different courses in both aspects of traditional learning systems or distance learning, the achievement test is the cornerstone of the educational process as the source of decision making on the level of success or failure to achieve objectives.

Therefore, the credibility of the evaluation provides the educational program with the basics of success as success depends on the effectiveness of evaluation methods, the procedures and the extent to which the evaluation process depends on the basic criteria, the principles that should be taken into account and good testing. Since the validity and reliability of the test depends on the characteristics of its test items, it is important to be aware of the conditions that must be met in the optional items and tests content according to the good testing criteria.

Due to the lack of test quality, in terms of test preparation, ambiguity or low level of the discrimination power and lack of good standards, many studies need to be conducted.

**Problem statement:**
Because the tests are important for the educational process, and since the quality of tests is one of the requirements of academic accreditation, it is necessary to establish the rules and foundations to ensure the success of the tests of the highest possible quality, in the light of the standards of the National Commission for Academic Accreditation, which ensures the quality of tests and in an attempt to contribute to solve the
problems related to the tests of achievement and how to prepare them to assess the students in terms of design of the test questions. In the light of the overall quality standards, this participation in this study, where the researchers identified the problem of research in the following question:

**What is the quality of the questions of achievement tests in university education in the light of quality standards?**

This main question can be branched into the following sub questions:

- How diverse are the test items and contents in the Faculty of Education in Afif?
- How much do the staff members at Faculty of education in Afif are aware of the criteria of good and quality testing?
- What is the level of commitment of staff members to standards to good and quality testing?
- What is the level of commitment of staff members to quality specifications in their test papers?

**The study objectives:**

The study aims to identify:

- The extent of the diversity of test items and contents in the faculty of education in Afif.
- The awareness of staff members in the faculty of education in Afif of good and quality testing standards.
- The level of commitment of staff members in the faculty of education in Afif to good and quality testing standards.
- Level of commitment of staff members of faculty of education in Afif to good and quality specifications in their test papers.
The importance of the study:
This lies in the importance of testing, which is characterized by scientific progress and the need to keep pace with the development of the education system through the development of tools to measure the students' performance according to the overall quality standards in order to match the outputs with the development of the ongoing labor market and away from the traditional pattern of tests that is based on conservation as the main estimation of academic achievement and as one of the indicators of determining the productive efficiency of the university. Educators also benefit from the development of measuring tools to achieve the quality of outputs.

Limitations of the Study:
- Objectivity: What extent is the quality of achievement tests in university education in light of the overall quality standards.
- Time: The first semester of the year 1435-1436 AH
- Spatial: Shaqra University - Faculty of Education for girls in Afif Governorate
- Terms: Standards - Quality - Achievement

Theoretical framework:
Firstly: The Concept of Total Quality
The process of improvement and development began from the beginning of creation, when the first person tried to secure his life, needs, requirements of food, shelter, clothing, weather and to protect himself against danger. Other factors began to invest the entire environment provided by the plant and animal products to make necessary needs, and did not stop at that, but gradually began to improve and regenerate what he produced, depending on the needs to reach a
better life; until he reached the tremendous science and technological development that was now achieved.

The key feature of the Muslim personality raised by Islam and Islamic education, is associated with the Muslim in his life and in his Islamic society to become a behavioral phenomenon. Workers must be proficient in their performance of work. Therefore, the development and improvement through the work is something realistic and recognized, as it is the path of success through creativity in work, and since we are called (read nation). Our attention to science and our responsibilities are very important. This received great attention by the government and officials of education in order to ensure that our graduated students achieved the educational and scientific targeted objectives and they can achieve our society distinctive morals with enough science and knowledge through quality.

The concept of quality and improvement of work in general is not new to Muslims, Islam urges us to master the work and calls for improvement, quality, and mastery in work and make those who improve the best work in the Koran and a lot of the Prophet's Hadeeth, such as saying (God Almighty) (Al-Kahf 30). It is a saheeh hadith that urges the improvement of the work, as the Prophet (peace and blessings of Allah be upon him) said: "Allah will be pleased with those who try to do their work in a perfect way" Narrated by Muslim. Al-Bayhaqi found that many of these principles were taken care of by Islam and worked to consolidate them, including but not limited to love of work and mastery, consultation, auditing and accountability, sincerity and self-monitoring, work teams, cooperation and solidarity among the members of society.
Quality in its Islamic disciplines look forward to elevate the student morality, scientifically and socially so that the benefit in society will be in accordance with the quality controls. Angeline and others (2006,6) discuss the quality in lightened of life skills (Life Skills encompasses social attitudes, basic knowledge and practical skills). It includes broader than vocational skills, practical skills and knowledge that lay the foundation for children to be economically productive when they enter the world of work. „New curricular areas or cross-cutting themes, such as peace education, health or education for sustainability, can be viewed as focusing on the relatively neglected pillars of learning to live together and learning to do.

**Secondly: The achievement tests:**

Tests are measured to detect the impact of learning or special training. This term is called upon on all types of tests that the teacher prepares from the reality of the learning materials that the student has already studied (HadiMishaan and Yatham Ismail, 2010, 153).

The achievement tests are one of the most important tools used to measure students' performance in courses at all levels of study. The most important feature of these tests is that they are made by faculty members because they prepare them themselves. In order to use these tests, it must be noted that each type has specific conditions and terms, but at the same time contain negatives, which necessitate the use of different types.

Achievement testing means the tool that is used to measure knowledge, understanding and skill in a particular subject or set of materials or is the test that is prepared by the subject teacher. Achievement is also understood as learning goals of the students, following
their study of a particular subject (Abdul Rahman Abdullah, 2011).

**Classification of Achievement Tests:**

Achievement tests have generally been categorized as single-subject tests, survey batteries, or diagnostic tests and further dichotomized as group-or individually administered tests. Reference to the Ninth Mental Measurement Yearbook (Mitchell, 1985) they have been categorized as (a) group administered, (b) individually administered, and (c) modality-specific tests of achievements, which can be either group or individually administered.

Furthermore it has been classified according to its function as:

- **Surveying:** It determines the level of the general performance of students.
- **Level of the tests:** It determines the depth of student understanding of the specific educational material.
- **Diagnostic tests:** These are used to determine the strengths and weaknesses of a specific student or group of students in a specific subject or set of topics.

**Classification based on method of preparation:**

- **The standardized tests:** These tests, which were prepared by a group of specialists and tested on a group of students after the amendment and determined the procedures and methods of estimating the results and external circumstances in which the appropriate time to conduct is determined (EzzedineAlmjzob, 2010, 17).
- **Unchecked tests:** These are prepared by the teacher and used on a limited scale such as the quarterly tests.
Classification is based on the method in which the test is conducted. It is divided into:

- Oral tests: tests that require the individual to give a verbal answer to a particular question.
- Written tests: The student answers the questions in writing and is divided into:
  - Objective tests: which are not affected by the score obtained by the student and are corrected by a person or a teacher.
  - Self-assessment tests: in which the student's score is affected by the correction of a person or a teacher.

In addition to the analytical and objective tests, there are performance tests that are concerned with evaluating the practical behavior in scientific subjects and standard tests in which each question contains a standard criterion or a general level of measurement based on the competencies of the knowledge, abilities and skills of the students in which the performance of students is not compared to each other but comparing it to a specific general standard. (Mandour Abdel Salam, 200 AD, 114-115)

Objectives of the achievement test:

In general, achievement tests are used to make decisions which may involve instructions, guidance, or administrative issues. For example, what is the efficacy of a particular method of instruction? What are the specific outcomes of learning? Is there a need for remediation? Are grading practices accurate? Is the curriculum responsive to the acquisition of basic and specific academic skills? Is counseling appropriate for any given student? Is appropriate placement a concern? Thus, the breadth of the assessment will be predicated upon the rationale for the use of particular achievement
measures. Table 7.2 illustrates the types of questions or problems that may be addressed and the expected benefits to be derived from the testing process (Lynda and Gregory, 2006)

Achievement tests aim to achieve several skills such as:
- Determining and assessing the student knowledge.
- Finding ways to help the learner to reach the possible academic achievement.
- Discovering different mental and emotional health and tracking the growth of the learner.
- Knowing the appropriateness of the curricula to the levels of the students.
- Instructing learners in terms of the type of specialization appropriate to their mental abilities for professions and appropriate work.
- Helping parents to know the levels of their children. (HadiMishaan and Yatham Ismail, 2010, 155)

The importance of the achievement tests:
There is an agreement among educators that the tests of achievement are of great importance that can be represented in the following:
- To determine individual differences between learners.
- To know the current educational level of the learners and what they can do.
- To stimulate learners' education and create a spirit of competition.
- To shift students to a higher semester.
- To evaluate the teacher.
- To standardize tests

Principles of the achievement tests:
- A thorough and systematic examination of the scientific subject of the test.
• Knowledge of the basics of the material and related materials for understanding.
• Determining the educational objectives of the material precisely to develop the test.
• Determining the relative weight of each goal.
• Translating educational objectives into procedural objectives.

**Previous studies:**

Abu Zina (2001) aimed to recognize the reality of school tests set by teachers to measure the collection and education of their students in the study of mathematics by analyzing a sample of these tests and stand on the procedures and methods used in the preparation of the study. He concluded that the development of tools and means of measuring and the introduction of new methods are in accordance with the specifications table reflects the relative weight of the subjects. Mahtoy Drash Mohammed bin Ali Ibn Mamedalsubha 2001 (build a test reference for measuring mathematical skills in the engineering concepts for the primary stage in Mecca government schools) and it was the main goal of the research. It was found out the extent of mastery of primary school students in Mecca sports competencies in engineering concepts: "visual skill, descriptive skill, logical skill, the skill of drawing". Somora (2002), study aims to seek quality legalization measure classroom tests in Los Angeles schools. The study sample consisted of 181 teachers selected randomly. The results showed the average approval among respondents who have registered answers at an average rate of efficiency. The study also recorded that high school students showed the highest levels in language tests in Qara Of the Stanford test completion rate after the scientific
background and past achievements. The study of Ghassani(2006) aimed to determine the basis for developing written tests in social studies and to know the reality of the written tests in the preparatory schools in Oman and to determine the criteria necessary for the development of written tests and the availability of standards in the written tests. (181) students participated in the study. The study concluded that the tests are designed to achieve the goals of social studies and that they develop students' abilities and measure the higher levels of thinking.

**Methodology:**

The researchers followed the analytical descriptive approach.

**The community of the study:**

Staff members of the Faculty of education are female students in Afif who are 72 staff members, with different academic degrees, academic specialization, different number of teaching performance training courses.

**The sample of the study:**

A random sample of 46 staff members of the Faculty of Education -Afif was selected.

**Description of the study Sample:**

Table (1) :Shows the sample of the study by academic degree.

| Academic degree          | Number | Percentage |
|--------------------------|--------|------------|
| Teaching assistant       | 3      | 6.5        |
| Lecturer                 | 20     | 43.5       |
| Assistant professor      | 21     | 45.7       |
| Associate professor      | 2      | 4.3        |
| Total                    | 46     | 100        |
Fig (1): Shows the sample of the study by academic degree

From the data of the table above we noted that the majority of faculty members are assistant professors, followed by the rank of lecturer and the lowest degree in the post of associate professor and the post of lecturer.

Table (2): Shows the members of the study sample by specialization

| Specialization     | Number | Percentage |
|--------------------|--------|------------|
| Educational        | 12     | 26.1       |
| Not educational    | 34     | 73.9       |
| Total              | 46     | 100        |

Fig (2): Shows the percentage of the study sample by specialization
As shown from the data at the table above, we noted that the majority of the faculty members and 73% of their Specialization from colleges other than faculty of education.

**Table (3): Shows the sample of the study according to year of experience**

| Teaching experience | Number | Percentage |
|---------------------|--------|------------|
| Less than 5 years   | 3      | 6.5%       |
| 5-10 years          | 20     | 43.5%      |
| More than 10 years  | 21     | 45.7%      |
| Total               | 100    |            |

**Fig (3): Shows the percentage of the sample according to year of experience**

From the table above we noted that the experience of the majority of faculty members is more than 10 years.

**Table (4): Shows the sample of the study who are trained to perform teaching efficiently.**

| Number of trained staff members | Number | Percentage% |
|---------------------------------|--------|-------------|
| Trained                         | 40     | 87%         |
| Not trained                     | 6      | 13%         |
| Total                           | 46     | 100%        |
Fig(4): Shows the percentage of the sample who are trained to perform teaching efficiently.

From the table above we noted that the majority of faculty members are trained to perform teaching efficiently.

Tools of the study:
A questionnaire designed by the researchers to gather staff members opinions about the extent of the achievement tests quality in university in light of the total quality standards. The questionnaire included:

- Part one: general data of the study sample which contains: academic degrees, academic specialization, training to perform teaching efficiently and year of experience.
- Part two: the extent of the variety of test items and question types.
- Part three: the extent of awareness of the standards of good testing.
- Part four: the level of adherence to the standards of good testing.
- Part five: the extent of adherence to the specifications of the exam paper.

Firstly, the standards and characteristics of the questionnaire:
The researchers applied the questionnaire which consists of four parts with total of (34) items, the
second part consists of (10) items, and (8) items for the third, fourth and fifth parts on an initial sample of (20) faculty members randomly selected from the study community.

The internal validity and reliability of the questionnaire items was calculated and the overall reliability of the whole were confirmed.

To find out the accuracy of the reliability of questionnaire items with the total degree of the current study community, the researchers used (Pearson correlation coefficient) to calculate the correlation between the scores of each item with the total score of the sub-scale under which the item concerned is located. The following table shows the results of this procedure.

**Secondly: Validity and Reliability:**

In order to determine the reliability of the overall scale in its final form it was applied to a sample of (24) members pre-actual use and then the reliability was calculated using (Alpha Kronbach formula) to the preliminary sample data. The results of this procedure are shown in the following table:

| TableNo. (5) shows the results of reliability and validity. |
|-----------------------------------------------|-----------------|-----------------|
| parts                                         | Number of Item | Factor          |
|                                              |                | Alpha | Self-Validity |
| The extent of the variety of test items and questions types. | 10              | 0.875 | 0.922         |
| The extent of awareness with the standards of good testing. | 8               | 0.935 | 0.966         |
| The level of adherence to the standards of good testing. | 8               | 0.947 | 0.937         |
| The extent of adherence to the specifications of the exam paper. | 8               | 0.966 | 0.928         |
From the table above we noticed that the tool has high reliability and validity of each of the tool items.

|                           | Overall reliability | Overall validity  |
|---------------------------|---------------------|-------------------|
|                           | 0.93                | 0.96              |

**Statistical analysis and results**

Statistical Methods: After the data was collected it was analyzed using SPSS program, several formulas were used to interpret results such as:

- Pearson Correlation, to determine the internal consistency between items and their related main and sub parts of the internal consistency of the study questionnaire.
- Alpha-Cronbach stability coefficient to determine the stability of the study questionnaire.
- Frequency and percentages to describe the sample of the study and the distribution of opinions between the study sample according to the variables.
- The arithmetic mean and the standard deviation of the study questionnaire.
- (ANOVA) to calculate the effect of the initial data at the study variables.

**Table No. (6): The frequencies and percentages of the opinion of study sample about the variety of test items and questions types.**

| Question type                              | Frequency | Percentage |
|--------------------------------------------|-----------|------------|
| Short assay (explain)                      | 17        | 11.7%      |
| Long essay (write about)                   | 9         | 6.2%       |
| Fill in the blanks                          | 18        | 12.4%      |
| Find the terminology                        | 17        | 11.7%      |
| Yes or No                                   | 19        | 13.1%      |
| Multiple-choice                             | 16        | 11.3%      |
| Matching                                   | 5         | 3.4%       |
| Questions of memorizing and reciting       | 17        | 11.7%      |
| Draw or explain on the drawing             | 15        | 10.3%      |
| Solving equations                           | 12        | 8.2%       |
| Total                                      | 145       | 100%       |

From the data in the table above we notedd that most question type used by the sample of the study
were (Yes or No questions ) with (13.1%) and then (fill in the blanks) with ( 12.4% ) and the least applied were the long essay with ( 6.2%) and matching questions with( 3.4%).

Table No. (7): Shows the frequencies, percentages, mean and standard deviation of the extent of awareness with the standards of good testing.

| statements | Scale responses | Average | Standard deviation | The result |
|------------|----------------|---------|--------------------|------------|
|            | correct         | True to some extent | Incorrect |             |       |
|            | frequencies | percentage | frequencies | percentage | frequencies | Percentage |       |       |
| 1- Comprehensive: means that the test covers all elements of the content and includes all domains of objectives as much as possible | 46 | 100 | 0 | 0 | 0 | 0 | 1 | 000 | Correct |
| 2- Objectivity: means the scores to which equally competent users get the same results. | 37 | 80.4 | 8 | 17.4 | 1 | 22 | 1.22 | 0.467 | Correct |
| 3- Discrimination is the ability of the test to highlight individual differences among students | 38 | 82.6 | 8 | 17.4 | 0 | 0 | 1.17 | 0.383 | Correct |
| 4- Normality: distribution of the grades and means the distribution of grades as desired by the examiner | 23 | 50 | 13 | 28.3 | 10 | 21.7 | 1.72 | 0.807 | True to some extent |
| 5- Justice means taking into consideration the table of specifications and the relative weight of the units of study when developing the test items. | 31 | 67.4 | 15 | 32.6 | 0 | 0 | 1.33 | 0.474 | Correct |
| 6- Validity: means intended to measure what it had been designed to measure. | 37 | 78.3 | 9 | 19.6 | 1 | 22 | 1.24 | 0.480 | Correct |
| 7- Reliability: means that we will have the same results if we re-test students at the same conditions | 29 | 63 | 13 | 28.3 | 4 | 8.7 | 1.46 | 0.657 | Correct |
| 8- Applicability means the possibility of applying a test according to a specified time | 33 | 71.7 | 12 | 26.1 | 1 | 22 | 1.30 | 0.511 | Correct |
From the data in the above table we noted that the sample failed to know the criteria of (2) the objectivity, (4) the natural distribution, (6) the validity and (8) the applicability, the majority chose the response (true) 80.4 - 50 - 78.3 - 71.7 with an average of 1.22 - 1.72 - 1.24 - 1.30. The responses were 50% correctly and this confirms that their knowledge about good testing criteria is average.

Table (8): Shows the frequencies, percentages, arithmetic mean, and standard deviation of the third axis expressions.

| statements                                                                 | Scale responses |          |          |          |          |          | The result   |
|----------------------------------------------------------------------------|-----------------|----------|----------|----------|----------|----------|--------------|
| The extent to which the test items cover the material.                    |                 | correct  | True to some extent | Incorrect | Average | standard deviation |            |
|                                                                            |                 | frequencies | percentages | frequencies | percentages | frequencies | Percentage |                |
|                                                                            |                 | 46        | 100       | 0         | 0         | 0         | 1           | 000         | Correct       |
| The extent to which the objectives of the subject have been achieved through testing | 37              | 80.4      | 8         | 17.4      | 1         | 22        | 1.22        | 0.467       | Correct       |
| The validity of the test in measuring what was set for it                 | 38              | 82.6      | 8         | 17.4      | 0         | 0         | 1.17        | 0.383       | Correct       |
| The variety of test items from measurement to memorization, understanding, application and analysis | 23              | 50        | 13        | 28.3      | 10        | 21.7      | 1.72        | 0.807       | True to some extent |
| The degree of accuracy and impartiality in correcting the test             | 31              | 67.4      | 15        | 32.6      | 0         | 0         | 1.33        | 0.474       | Correct       |
| The extent of the distribution of test scores according to the ease and difficulty of the paragraph | 36              | 78.3      | 9         | 19.6      | 1         | 22        | 1.24        | 0.480       | Correct       |
| The extent of the test ability to distinguish differences between the achievement scores and can be represented by a natural curve | 29              | 63        | 13        | 28.3      | 4         | 8.7       | 1.46        | 0.657       | Correct       |
| The extent to which the question paragraphs and their degrees are related to the effort to design it | 33              | 71.7      | 12        | 26.1      | 1         | 22        | 1.30        | 0.511       | Correct       |

From the above table data we noted that the highest frequencies response of the study sample was (high). The mean of all the expressions is convergent as well as
the standard deviation. This confirms that the respondents' compliance with the test criteria is high, even though they do not understand the meaning of some criteria.

Table No. (9)" Shows the frequencies, percentages, arithmetical mean and standard deviation of the fourth axis expressions.

| Statements                                                                 | Scale responses | Avera ge | standard deviati on | The result |
|---------------------------------------------------------------------------|-----------------|---------|---------------------|------------|
| 1-The question sheet includes the students' personal information, the course information, the test instructions, and the grade table | Repetition | 40 87 4 87 2 4.3 | 117 | 0.486 | I agree |
| 2-The test paper is clearly printed and will design.                      | Repetition | 40 87 6 13 0 0.0 | 1.13 | 0.341 | I agree |
| 3- Formulation of the questions and their integrity in a grammatical manner and free of spelling mistakes | Repetition | 42 91.3 4 8.7 0 0.0 | 1.09 | 0.285 | I agree |
| 4- Signing the question sheet from the instructor.                        | Repetition | 31 67.4 14 30.4 1 2.2 | 1.35 | 0.526 | I agree |
| 5- Taking into consideration not pressing the questions in one sheet.      | Repetition | 38 82.6 6 13 2 4.3 | 1.22 | 0.513 | I agree |
| 6. The adequacy of the questions to measure the expected learning outcomes of the course and is committed to describing the course and its objectives | Repetition | 38 82.6 6 13 2 4.3 | 1.22 | 0.513 | I agree |
| 7- Distribution of grades to the main and secondary test questions         | Repetition | 36 78.3 6 13 4 8.7 | 1.30 | 0.628 | I agree |
| 8- Reinforcing the question sheet for students such as using phrases like (with my wishes for success) | Repetition | 39 48.8 6 13 1 22 | 1.17 | 0.437 | I agree |

From the above table data we noted that the majority of the study sample respondes with ( agree) so it took the highest frequency and mean. Since the mean of all expressions is convergent as well as the standard deviation that confirms that they adhere to the quality specifications in preparing the test paper.

Conclusions:

The study reached several results:

- The tests contents which are prepared by the staff members of the Faculty of Education in Afif vary
properly, and most questions types used by the sample of the study were (Yes or No questions) then (fill the blanks) then the long essay questions and the least applied were the matching questions.

- The awareness of faculty staff members at the Faculty of Education Afif in terms of the standards of good and quality testing standards was average.

- The level of commitment to the quality testing standards of faculty staff members in the Faculty of Education Afif was high.

- The level of commitment of faculty staff members in the Faculty of Education Afif to quality specifications in their test papers was good.

**Recommendations**

In the light of the results and findings of the study the researchers recommends the following:

- To provide training courses for faculty staff members who hold qualifications other than the faculties of education in measurement and evaluation criteria in general and in the achievement tests in particular.

- It is necessary to raise the staff members awareness of educational goals and how to formulate different types and questions and translate them to behavioural goals which can be measured.

- To training teachers on teaching and evaluation methods, since the achievement tests are important and essential to assess different aspects of learning skills. The faculty staff members must be familiar with the basics of teaching methods and methods of evaluation.
References:
- The Holy Quran.
- Abdul Rahman Ahmad Abdullah, Amani Abdulrahman Makkawi, Noha Ibrahim Al-Khalil 2011, Methods of Measurement, Evaluation and Statistics, Sudan University Press for Science and Technology, Sudan
- Angeline M. Barrett, Rita Chawla-Duggan, John Lowe, Jutta Nike, Eugenia (2006): THE Concept of quality in education: a review of the ‘international’ literature on the concept of quality in education, EdQual Working Paper No. 3, University of Bristol, University of Bath, P.6
- Ezz Eddine Abdel Rahim (2010) Measurement and Evaluation, Sudan University of Science and Technology, Sudan
- Farid Abouzineh (2001), Development of tools and tools for students in Mathematics, Journal of Research Center, University of Qatar, Year (10th) of the No. (19) pp. 107-107
- Hadi Mishaan Rabee and Yatham Ismail Ahmad (2010) Measurement and Evaluation in Education, Dar Zahran Publishing and Distribution, Amman
- Handbook of Psychological Assessment (Third Edition), 2000 which available at: https://www.sciencedirect.com/topics/medicine-and-dentistry/achievement-test last visited 30 Dec 2019
- Lynda J. Katz, Gregory T. Slomka (2006): http://cachescan.bcub.ro/e-book/E1/580683/149-201.pdf
- Mohammad Ali Humaid Al-Subhi (2001), Building of the test of the auditor of the educational system in the engineering concepts of the primary stage of the Ministry of Education, Ministry of Higher Education, University of Makkah.
- Mohammed Salim Alghassani (2006), Development of the tests in the field of social studies for the students of the preparatory stage in the Sultanate of Oman in light of the objectives of the teaching of the subject, the thesis of the Master of Science, University of Cairo.
- Mastomra (2002), Shalaa, Muhammad Al-Zahrani (2009), A Proposal for the Development of Tools for the Achievement of
Students and Comprehensive Quality Standards in the Ministry of Education and Education, University of Umm Al-Qura, Makkah
- Mandour Abdul Salam Fath Allah (200) Educational Calendar, First Edition, International Publishing, Saudi Arabia – Riyadh
- Ziad Birkat, Abdul Hadi Sbah (2007), The Examination of the Final Exam Questions at the Open University of Jerusalem for Academic Goals, Jerusalem University Journal for Research and Studies, No.9

***