

The Degree of the Practice of Ma'an Governorate Teachers for the Skills of Building Achievement Tests

Dr. Raed Omar Salah

Assistant Professor of Methods of Teaching in Mathematics

College of Educational Sciences

Al Hussein Bin Talal University

Dr. Habis saad Al-Zboon (Corresponding author)

Associate Professor of Assessment and Evaluation

College of Educational Sciences

Al Hussein Bin Talal University

Dr. Mansour Hamed AlTalhani

Assistant Professor of Curricula and Teaching Methods in Islamic Education

College of Educational Sciences

Al Hussein Bin Talal University

Received: November 17, 2019 Accepted: January 7, 2020 Published: February 5, 2020

doi:10.5296/ijld.v10i1.15845 URL: <https://doi.org/10.5296/ijld.v10i1.15845>

Abstract

This study aimed to identify the degree of practice of the teachers in Ma'an Governorate for the skills of building achievement tests. In order to achieve the objectives of the study, a questionnaire was developed consisting of (36) items divided into two domains. After confirming the psychometric properties of the questionnaire, it was applied on the study sample which consisted of (186) male and female teachers at Ma'an Governorate. The study sample was randomly selected.

The results of the study indicated that the degree of the practice of the teachers of Ma'an Governorate for the skills of building the achievement tests was average. They also indicated

that there were no statistically significant differences according to the gender variable, and there were statistically significance differences attributed to the variable of specialization in favor of the scientific specialties over the humanities. Furthermore, there were statistically significant differences attributed to the experience and qualifications variables in favor of the master's degree holders compared to those of bachelor's degrees. It was also found that the most important needs to build the achievement tests for the teachers of Ma'an Governorate were to provide special evidence for the building of the achievement tests, followed by conducting training sessions specialized in the building of tests for the preparation of new teachers.

Keywords: achievement tests, Ma'an teachers.

1. Introduction

Achievement tests are one of the most important measurement tools in the educational process since they are an integral part of the assessment process that has become important in the curriculum. They are also a component of the teaching and learning process. They provide us with indicators that show how successful the education process is. It also assists in reviewing all aspects of the education process with a view of developing it in the light of what the tests refer to (Metwally, Al-Tahawi, Qastanian, & Al-Rabai, 2005).

Moreover, achievement tests are not only one of the main methods of evaluation but also the most widely used in the school environment, which dates back to the first quarter of the 20th century, as some educators pointed out, including Al-Samara'i (1985) and Muncie (1998). Despite scientific progress in all fields, the teacher's preparation of the educational tests is still weak in such a way that the results of these methods are doubtful (Al-Zahrani, 2003).

The development of achievement tests is an important part of measurement and evaluation in any learning environment, and both teachers and supervisors are expected to have a certain level of mastery in building reliable and valid tests to adequately evaluate students' performance. Since many decisions concerning students are based on the latter's performance in different tests. Students may or may not graduate based on tests' results, which makes the development of a good achievement test a very important task for teachers (Simsek, 2016).

As school tests are an important evaluative tool in the educational system, and almost all other educational matters are dominated, we often hear parents saying that their children have scored high grades. Parents also declare that some of their children are distinguished because their grades are higher than their peers' and others are subject to study delays because their grades are lower than their peers' and so on. The educational evaluation is one of the main pillars of the educational process, and is an important part of its steps through which students are evaluated and the educational process is evaluated in all its aspects (Bloom, 1973).

Teachers rely on students' results in tests and other assessment methods that teachers use to make judgments that lead to important decisions related to different aspects of the educational process such as student success, modification of the teaching method, the distribution of students in groups, etc. The concept of evaluation involves making effective judgments about activities, individuals or programs, and in order to be accurate and objective, we rely on

quantitative data collected using different measurement and evaluation tools (Oudeh, 2010).

Therefore, the teacher must follow the scientific method of obtaining the data and information s/he wants about students, using appropriate evaluation tools, so that this information is accurate and appropriate. In order to do so, it is important that teachers should know the principles of measurement and educational evaluation and its skills to prepare tests that can be used efficiently in their evaluation processes (Melhem, 2002).

At the end of the teaching process, teachers conduct a summative assessment in which they conduct achievement tests and measure what has been achieved. These tests are prepared by teachers and determine the students' success or failure. This significant role of testing in the teaching process requires that teachers should know the concepts related to the building of school tests, which are concepts based on theories, facts, and information. Teachers must therefore be able to develop concepts necessary to build achievement tests as one of the pre-assessment measurement tools based on results obtained by many decisions. The importance of this study is therefore to identify the degree of practicing the skills of building the achievement tests that Ma'an Governorate teachers have.

1.1 The Problem of the Study

The purpose of this study is to identify the degree of practicing the skills of building the achievement tests that Ma'an Governorate teachers have. Achievement tests are one of the most important tools that teachers rely on in issuing judgments leading to important decisions related to different aspects of the educational process, such as student success, modifying the teaching methods, the distribution of students in groups, etc. The problem of the present study was formed through the direct contact of the researchers with teachers in the master's and diploma program at the Faculty of Education. Through teaching some courses such as the school exam, it is shown that there are many problems that teachers face in building the achievement tests according to the criteria of a good test. Training teachers on building achievement tests helps them improve decision-making that could be based on the results of the tests, particularly in the absence of training programs for teachers on the building of achievement tests. In order to identify the degree of practicing the skills of building the achievement tests that Ma'an Governorate teachers have from the teachers' point of view, this study therefore attempts to answer the following questions:

1. What extent of knowledge do Ma'an Governorate teachers have concerning practicing the skills of building achievement tests?
2. Are there any statistically significant differences at the indicative level of ($\alpha=0.05$) in the extent of knowledge of Ma'an Governorate teachers concerning practicing the skills of building achievement tests, which is attributed to the gender, qualification and specialization, and experience variables?
3. What are the most important requirements for the building of achievement tests for Ma'an Governorate teachers?

1.2 The Study Terms and Procedural Definitions

The skills of building tests: The ability of teachers to perform all procedures and tasks of building achievement tests so that all specifications of a good test are considered, and this skill is measured procedurally to the extent that the teacher gets a test based on the achievement tests prepared by the researchers.

The achievement test: An organized procedure for measuring a sample of an individual's behavior (Chase, 1977), measured procedurally by the score which the responder gets through the designed test.

The determinants of the study:

The accuracy of the results is determined by the extent to which the tool developed to achieve its objectives through indicators of validity and reliability.

Moreover, the study was applied during the second semester 2018-2019 on Ma'an Governorate male and female teachers.

2. Theoretical Framework and Literature Review

Tests are one of the most important ways to measure and assess the students' abilities so as to determine their level of achievement, to know to which extent the behavioral goals or learning outcomes have been achieved and what the teacher provides of various educational activities that help raise students' competencies. Therefore, educational supervisors are keen to ensure that tests are highly competent in the measurement and evaluation process (Al-Najjar, 2010).

Since the achievement tests are important, the preparation of any test must be taken seriously, and enough time must be allocated for its preparation. The teacher, while preparing the test, should consider the importance of preparing this tool very well since s/he is considering a tool that governs his/her students' level. The teacher, in considering the psychological effects of this tool and the social dimensions of the student's personality, must be accurate and cautious in preparing the test. Achievement tests have obtained the educationalists' interest in the development of good tests in order to produce good results that can be adopted in educational decision-making (Ministry of Education, 2017).

Achievement tests are considered one of the main components of the educational process, where the teacher has the responsibility of determining the future of his/her students. The teacher who has the ability and skill to teach his/her students must also have the ability to test them through his/her ability to develop achievement tests for the courses s/he teaches (Abdel Hadi, 2001).

A good achievement test must have several basic characteristics to be entrusted and benefited from its results, including, as mentioned by (Abu Libda 1985; Oudeh, 2010), validity and comprehensiveness.

Validity means that the test measures what is set to measure, whereas comprehensiveness means that the test includes all teaching goals to be measured taking into account the level of knowledge and the relative weight it represents.

3. Reliability and Objectivity

The reliable test is the test whose results are close or have the same results if the test is applied more than once in similar circumstances, which means that every examinee maintains almost the same position in his/her group. Objectivity means that the examiner is unbiased in all steps of preparing and marking the test.

As Adas (1999) pointed out that a good achievement test must have a number of characteristics and principles in addition to its validity and reliability, including:

- Measuring specific learning outcomes, be consistent with teaching goals.
- Covering a representative sample of learning outcomes and the teaching material.
- Containing the quality of questions that are most appropriate for measuring desired learning outcome.
- Being developed to suit the purposes for which the results will be used.
- Having the possible optimal consistency and be cautious in interpreting the test results.
- Aiming at improving students' learning.
- Having the validity of the wording and clarity of the test items and being familiar in terms of form to examinees.

Many studies tackled this issue, including the study of (Boothroged, *et al*, 1992). This study aimed at identifying the nature and extent of the training on measurement received by teachers and revealing the skills and knowledge they possess and the characteristics of the teacher that relate to the knowledge of measurement. The sample of the study was formed of (41) Grades 7-8 Math and Science teachers. A multiple-choice test of 65 items was administered. The results of the study showed that teachers usually evaluate their students with tests that they develop themselves. The results also indicated that teachers have insufficient knowledge concerning measurement, owing to the lack of training.

Another study conducted by (Daniel & King, 1998) aimed to identify the knowledge of primary and secondary school teachers concerning educational tests, the principles of measurement and application of the various concepts of measurement and educational tests in the classroom assessment environment. The study also aimed at identifying differences in assessment strategies between primary school teachers and their peers in the secondary schools. 95 teachers were surveyed using the educational test and measurement and the use of assessment strategies at the primary and secondary schools and other demographic information. Generally, the study found that there was an inadequate knowledge concerning the concepts and basics of educational test among the teachers. Nevertheless, they always used what they know of measurement skills when assessing their students' learning and development. It was also found that there were few differences among the teachers working at the primary and secondary schools in learning and using evaluation practices.

Similarly, Hanno (2002) examined the extent of the use of the Arabic language teacher of the methods of evaluating primary school students in UNRWA schools in the district of Nablus Educational Zone. The population of the study comprised of all teachers who teach the Arabic language curriculum in the primary schools. The researcher developed a questionnaire

consisted of (81) items distributed into seven domains. The results of the study showed that the highest mean was on the scores concerning the extent of using the assessment methods in assessing students of Arabic language course in the domain of oral tests, listening tests, the written expression test and then speed test. Whereas the domain of the objective essay test scored the lowest means. The results of the study also indicated that the extent of the use of female teachers of the good assessment methods was better compared to male teachers, and that male teachers with a bachelor's degree applied better assessment methods than teachers with diplomas.

Boston and Carol's (2003) study summarized the teachers' practices in evaluating their students. The results of the study showed that many teachers had adopted tests that measured students recalling skill, and their grades were self-governing. It was not surprising that there were differences in student scores and achievements according to their region or state because of the way in which teachers assign marks on the tests and how they develop them.

The study of Chen (2007), at New Orleans University, aimed at revealing the types of assessment used to assess English learners as a foreign language among the different teaching methods. The sample of the study consisted of (12) secondary school students studying English as a foreign language at New Orleans University. The study tools were observation, interview, video recording, and a collection of student writings. The results of the study showed that the most commonly used assessment methods were essay questions, student portfolio, check list, and rating scale. Students expressed a desire to use more varied assessment methods such as discussion, peer-assessment, and interview.

In his study, Abu al-Haj (2010) aimed at identifying the extent to which upper primary stage English language teachers at the capital city of Jordan, Amman, know and use the alternative assessment methods. The sample of the study consisted of (150) English language male and female teachers. The researcher developed two research tools for this study (test and questionnaire). The study arrived at several results, including that the degree of knowledge of alternative assessment methods among English teachers in the upper primary stage was medium, and that their degree of use was high. The results also showed that there is a positive, statistically significant relationship between the knowledge and use of English language teachers at the upper primary stage of alternative assessment methods.

Caliskanm (2011) conducted a study about the primary stage teachers' opinion in general tests. To achieve the objectives of the study, it was applied to (60) teachers in the various fields of teaching in Ankara, focusing on the personal essay. The study aimed to answer a series of research questions related to the detection of errors in designing general tests in order to be avoided and to develop appropriate solutions, and to identify the positive aspects to enhance them.

Saada's (2012) study aimed at building a program to develop the competencies of physical education teachers in the domain of measurement and evaluation based on their training needs. Consequently, the level of competencies and skills of physical education teachers in the domains of measurement and evaluation was assessed in addition to the diagnosis of their weaknesses and strengths. To achieve the objectives of the study, a list was established to

survey the needs of physical education teachers concerning measurement and evaluation. The study also assessed the concepts, terminology, measurement and evaluation skills among the physical education teachers. The results of the study indicated that there were several factors related to measurement and evaluation that physical education teachers do not consider while performing their professional roles, the most prominent of which was the interdependence and complementarity in the knowledge about the evaluation tool. The results also showed that the practices of measurement and evaluation related to school administration and parents were limited.

The main obstacles facing teachers in practicing their professional roles in measurement and evaluation were those related to the instructions of the ministry. The second important obstacle was the problems related to the average number of students in class and the difficulty it represents in applying measurement and assessment procedures.

Moreover, the results of the study showed that there were statistically significant differences between the average estimates of teachers due to the variable of academic courses in measurement and evaluation concerning the formulation and identification of educational goals for teachers who have received courses related to assessment and evaluation.

The results of the study also showed that there were no statistically significant differences between the average estimates of teachers due to the variable of training workshops on measurement and evaluation. They also showed that there were statistically significant differences between the average estimates of teachers concerning their training needs due to the scientific qualification variable and to the benefit of the teachers of higher degrees.

Otom (2014) conducted a study aimed at identifying the compatibility of the achievement tests prepared by male and female teachers of schools at the Directorate of Education in Jerash Governorate according to the criteria of a good test. The study sample was formed of (116) items of an achievement test on the science, physics, chemistry, biology and earth sciences courses. The study showed that the achievement tests prepared by teachers were in line with the criteria of a good test and that there were statistically significant differences between male and female teachers in the extent to which the test matches the criteria of a good test in favor of female teachers.

Al-Omari (2015) also conducted a study aimed at determining the extent of knowledge which science teachers of the primary stage have of the principles of building achievement tests, and their ability to use and apply these principles in their working life. The study sample consisted of (103) male and female teachers. To achieve the objectives of the study, the researcher developed a tool consisting of (55) items distributed on five domains. Among the most important results of the study is that the knowledge of the principles of building the achievement test among the teachers of the primary stage, at the Directorate of Education in the Northern Mazar District, was average. The results of the study also showed that there were statistically significant differences in the principles of building achievement tests attributed to the gender variable and in favor of the female teachers. Moreover, the results showed statistically significant differences in the principles of building achievement tests attributed to the variable of academic qualifications and to the favor of the holders of bachelor's degrees in

preparation and planning.

The aim of Al-Kharabsheh (2017) study was to reveal the extent to which science primary stage teachers in the schools of the Ministry of Education in Al-Mafraq Governorate applied the skills of designing and building achievement tests. In order to achieve the objectives of the study, the researcher developed a questionnaire consisting of (31) items. One of the most important results of the study was that the extent to which science primary stage teachers applied the skills of building achievement tests was high. The results also indicated that there were no statistically significant differences between the average estimates of the teacher for their application of the building of the achievement tests due to the variables of gender, academic qualifications, training courses, and the existence of statistically significant and valid differences in favor of those of (5-10 years) of experience.

The study of Alhusban (2018) aimed at evaluating the achievement tests for the subject of vocational education at the primary stage in light of the criteria for the development of tests and the difficulties faced by vocational education teachers in the process of test preparation. In order to achieve the objectives of the study, two tools were developed. The sample of study consisted of 300 test papers for vocational education in Al-Zarqa Governorate. The difficulties questionnaire was applied to (100) male and female teachers whose test papers were included in the study sample. One of the most important findings of the study was that the criteria for a good test were average, and there were statistically significant differences depending on the gender variable concerning the efficiency of teachers in the final designing of the achievement tests in favor of the female teachers. The results also indicated that there were statistically significant differences in the average difficulties teachers face in building and building tests according to the gender variable in favor of the male teachers, and there were no statistically significant differences attributed to the variables of (academic qualifications, experience, training courses).

4. Methods and Procedures

4.1 The Study Population

The study population consisted of 755 male and female teachers at the public schools of Ma'an Governorate, (271) male teachers and (484) female teachers.

4.2 The Study Sample

The study sample, which was selected based on the simple random sampling method, consisted of (186) male and female teachers at Ma'an secondary public school at Ma'an Governorate.

4.3 The Study Tool

In order to achieve the objectives of this study, the relevant literature review which is related to the skills of the building of the achievement tests and the measures used was examined. The study tool which was developed consisted of two parts. The first included the data related to respondents whereas the second part included two domains: the first is related to the degree of practice of the teachers of Ma'an Governorate for the skills of building the achievement tests and consisted of (23) items. The second domain consisted of (10) items that measured the

needs for the development of achievement tests, after reviewing the previous local studies in this field.

4.4 Tool Validity

First: The validity of the reviewers: To verify the validity of the questionnaire, to ensure that it conforms to the goal for which it was prepared, and to assess its apparent validity and give their opinions on the validity, comprehensiveness and integration of its items, the questionnaire was presented to a group of faculty members in the Jordanian universities, specialists in the field of measurement and evaluation, and educational supervisors. reviewers were asked to delete inappropriate items or modify or move items from one domain to another or suggest any other proposals they deemed appropriate. Their comments were considered, where two items were added to the first domain. Some others were modified based on the ratio of agreement between the reviewers, as the percentage of agreement among them was (85%) on the retention of the item, and that the items that do not achieve this test have been deleted, either due to the lack of clarity or relevancy. Thus, the final questionnaire consisted of (35) items distributed on two main domains, 25 items for the first domain and 11 for the second one.

Second: The validity of internal consistency: The validity of the internal consistency of the questionnaire was verified after being applied to the members of the pilot sample of (33) male and female teachers outside the study sample. Pearson correlation coefficient was calculated between the scores of each domain of the questionnaire and its total score. The correlation coefficient of the first domain with the total score was equal to (0.83) and significant at ($\alpha = 0.01$), whereas the correlation coefficient of the second domain with the total score was equal to (0.89) and significant at ($\alpha = 0.01$). This shows that the two domains are statistically significant at the significance level ($\alpha = 0.01$). This confirms that the questionnaire has a high degree of validity and internal consistency.

Pearson correlation coefficient was also calculated between each item of the questionnaire and the total score of the domain to which the item belongs. It was found that all the items are correlated to the overall score of the domain and their correlation was statistically significant, which shows the validity of the internal consistency of the items of the questionnaire.

4.5 Tool Reliability

The reliability of the tool was estimated after applying it to the members of the pilot sample consisting of (33) male and female teachers, using the split-half method and the Cronbach Alpha method. To eliminate the effect of the split-half, the Spearman-Brown equation was calculated. The reliability coefficient based on the split-half method was (0.84) and on the Cronbach Alpha method (0.88). This indicates that the questionnaire enjoyed a high degree of reliability, which reassures the researchers to apply it to the study sample.

5. Study Results and Discussion

To answer the first question, "What is the degree of practice of teachers of Ma'an Governorate for the skills of building achievement tests?", the mean, standard deviation (SD) of the study tool items, and the reviewers' opinions, concerning distributing the respondents' responses into

three levels where the level is judged based on the below criterion, were relied upon:

Category Length = Range divided by the number of categories; accordingly, the following categories were adopted:

1. Less than 2.33 is low
2. 2.34 to less than 3.67 is medium
3. 3.68 and above is high

Table 1 illustrates this:

Table 1. Mean and standard deviation of the study sample answers for all study tool items arranged descending according to the mean

S. No.	Item No.	Item	Mean	SD	Degree
1.	25	I clearly mark each item	4.59	5.06	High
2.	11	I consider the abilities of the students to suit the time of answering the test questions	4.24	0.75	High
3.	7	I formulate test items considering individual differences among students	4.22	0.77	High
4.	6	I vary between the objective and essay questions	4.22	0.83	High
5.	2	I develop test items that measure the educational goals I seek to achieve	4.13	0.75	High
6.	1	I analyze the content of the course I teach	4.12	0.78	High
7.	10	I develop the test format appropriately	4.05	0.92	High
8.	18	I keep a question bank for each of the courses I teach	3.97	1.05	High
9.	24	I create separate items, so that the answer of one item does not relate to the answer of another one	3.75	0.99	High
10.	4	I build the specifications table before preparing the achievement test	3.72	1.03	High
11.	8	I include instructions for each test on the first page of the test	3.71	1.23	High
12.	5	I write a set of questions initially while teaching the course	3.69	1.11	High
13.	13	I adopt a model answer to mark the answers of the	3.68	1.08	High

		respondents to the essay questions			
14.	3	I develop cognitive goals according to Bloom's levels	3.6	0.91	Average
15.	15	I identify the marking procedures for the examinees	3.59	0.92	Average
16.	23	I develop the test items so that each one contains only one idea.	3.54	1.07	Average
17.	12	I ensure that my test doesn't include elective questions	3.51	1.32	Average
18.	14	I use a model answer to mark the objective questions	3.47	1.11	Average
19.	17	I explain the results of the analysis of the test application	3.43	1.03	Average
20.	9	I use difficult alternatives in the test items	3.3	1.17	Average
21.	16	I statistically analyze the test results	3.27	1.09	Average
22.	22	I find the coefficients of validity and reliability of the test scores	2.26	1.07	Low
23.	21	I calculate the distinguishing coefficients for the test questions	2.27	1.15	Low
24.	20	I calculate the difficulty coefficients for the test questions	2.23	1.12	Low
25.	19	I use different statistical programs to analyze the results of tests such as SPSS	2.2	1.24	Low
Total Score			3.55	1.18	Average

It is clear from Table 1 that the degree of practice of teachers of Ma'an Governorate for the skills of building achievement tests was average with a mean of (3.55) and standard deviation of (1.18). It is also clear that the item representing the highest degree of knowledge of teachers for building skills of achievement tests was (25), which states "I clearly mark each item," with a mean of (4.59) and a high degree. The next high degree is represented in item (11), "I consider the abilities of the students to suit the time of answering the test questions," with a mean of (4.24) and a high degree, followed by item (7), "I formulate test items considering individual differences among students," with a mean of (4.22) and a high degree. It is noticed that the lowest items representing the lowest means were item (19), "I use different statistical programs to analyze the results of tests such as SPSS," with a mean of (2.2) and a low degree, followed by item (20), which reads "I calculate the difficulty coefficients for the test questions," with a mean of (2.23) and a low degree, followed by item (21) "I calculate the distinguishing coefficients for the test questions," with a mean of (2.27) and a low degree,

followed by item (22), "I find the coefficients of validity and reliability of the test scores" with a mean of (2.26) and a low degree. The next low degree is item (16), "I statistically analyze the test results" with a mean of (3.27) and an average degree.

This shows that the degree of knowledge of teachers was high in the formal aspects of the test, while the aspects related to the analysis of test results, whether total or detailed, were low as the extent of teachers to find the psychometric properties of the test (validity and reliability) and psychometric properties of the items (coefficients of difficulty and discrimination) were low. The teachers' analysis of the students' scores on the test was average, perhaps due to the low knowledge of the teachers to analyze the test results and the lack of enough experience and training to analyze the results of the students' scores on the tests.

To answer question 2, which reads "Are there any statistically significant differences at the indicative level of ($\alpha=0.05$) in the extent of knowledge of Ma'an Governorate teachers concerning practicing the skills of building achievement tests, which is attributed to the gender, qualification and specialization, and experience variables?" mean and standard deviations were used. Table 2 illustrates this:

Table 2. Mean and standard deviation of the level of knowledge of Ma'an Governorate teachers concerning practicing the skills of building achievement tests, which is attributed to the gender, qualification and specialization, and experience variables

Variable		Mean	SD
Gender	Male	89.4	15.4
	Female	91.6	14.1
Qualification	Bachelor	88.9	13.7
	Diploma	94.5	14.5
	Master	111.3	14.2
Specialization	Scientific	93.8	15.5
	Humane	87.9	13.2
Experience	Less than 15 year	88.5	14.1
	5-10	93.01	15.1
	10-15	89.2	15.5
	More than 15	92.4	10.2

It is clear from Table 2 that there are apparent differences between the means, and to find out whether the differences are statistically significant, the Univariate Analysis of Variance was

used, and Table 3 shows that.

Table 3. Univariate Analysis of Variance attributed to the gender, qualification and specialization, and experience variables

Source of variance	Sum of squares	DF	Means squares	F	Sig
Gender	193.04	1	193.04	1.09	0.297
Qualification	1957.9	1	1957.9	11.08	0.001
Specialization	4584.94	2	2292.5	12.97	0.000
Experience	1789.77	3	596.6	3.38	0.020
Error	31289.4	177	176.8		
Total	39406.4	184			

It is clear from the table that there are no statistically significant differences attributed to the gender variable. It is also clear that there are statistically significant differences attributed to the variable of specialization and for the benefit of scientific specializations over the humanities. This may be due to the nature of the specialization as the teachers of scientific disciplines could deal with special procedures in the building of achievement tests, which relies heavily on analysis and calculations. Furthermore, it shows that there are statistically significant differences attributed to the variables of qualifications and experience, and to know the significance of differences, Scheffe test was used, and table (4) shows that:

Table 4. Multiple comparisons attributed to the qualification and experience variables

Variables							
Qualification				Experience			
	Bachelor	Diploma	Master	Less than 5 years	5-10	10-15	More than 15 years
Bachelor		-5.6	*-22.3	Less than 5 years	-4.5	-0.7	-3.85
Diploma			*-16.7	5-10		3.8	-0.66
				10-15			-3.1

It is clear from table (4) and from the multiple comparisons that the differences were significant between the bachelor and master's degrees and in favor of the master. This could be attributed

to the reason that teachers with a master's degree might have studied courses related to building achievement tests. It is also clear that the differences due to the variable of experience as a result of multiple comparisons are not significant as the differences between the averages were simple.

To answer the third question, which reads: "What are the most important requirements for the building of achievement tests for Ma'an Governorate teachers?" the means and standard deviations of the items were used.

Table 5. Means and standard deviations of the responses of the study sample for all items of the study tool in descending order according to the mean

S. No.	Item No.	Item	Mean	SD	Degree
1.	6	Providing evidence for building achievement tests.	4.23	0.87	High
2.	3	Holding specialized training courses in building achievement tests to prepare new teachers	4.22	0.94	High
3.	4	Holding training courses for teachers on the preparation of achievement tests	4.14	0.94	High
4.	7	Organizing mutual visits between teachers within each school	4.13	0.90	High
5.	5	Holding training courses on statistical analysis of test results	4.10	0.95	High
6.	10	Holding training courses in the field of writing tests items of all kinds	4.07	0.85	High
7.	8	Organizing mutual visits between teachers in different schools	4.04	1.03	High
8.	2	Providing specialized supervisors in the field of measurement and evaluation in each directorate	4.02	0.91	High
9.	11	Holding training courses in the field of calculating the psychometric properties of the tests	3.93	1.10	High
10.	1	Enrollment in the General Diploma Program in Education	3.90	0.97	High
11.	9	Holding training courses in the field of preparing the questions bank	3.80	1.07	High
Total Score			4.10	1.18	Average

It is clear from the previous table that the most important needs for building achievement tests for teachers of Ma'an Governorate are as follows:

First, providing special evidence for building achievement tests. Second, holding specialized training courses in building achievement tests to prepare new teachers. Third, holding training courses for teachers on the preparation of achievement tests. Fourth, organizing mutual visits between teachers within each school. Fifth, holding training courses on statistical analysis of test results. This shows that the most important needs centered on holding training courses for teachers' development.

Recommendations:

1. Providing training courses for teachers related to the use of different statistical programs to analyze the results of tests such as SPSS
2. Training teachers on the steps of building achievement test and analysis of test results, and
3. Providing evidence for building achievement tests.

References

- Abdulahdi, N. (2001). *Measurement and educational evaluation and its use in the field of classroom teaching* (second ed.). Amman, Jordan: Dar Wael for publication and Distribution.
- Abu al-Haj, N. (2005). *The extent of knowledge and use of alternative evaluation methods by teachers of English language in Amman upper primary stage*. (Unpublished doctoral dissertation). University of Jordan, Amman, Jordan.
- Abu Libdeh, S. (1985). *Principles of psychological measurement and educational assessment*. University of Jordan, Amman: Jordan.
- Adas, A. (1999). *Teacher's guide in building tests* (Second ed.). Amman, Jordan.
- Alhusban, M. (2018). *Evaluation of achievement tests for vocational education for the primary stage in the light of the criteria of building tests*. (Unpublished doctoral dissertation). University of Jordan, Amman, Jordan.
- Al-Kharabsheh, A. (2017). *The extent of applying science teachers at the primary stage in the schools of the Ministry of Education at Al-Mafraq Governorate for the designing and building of achievement tests*. (Unpublished Master Thesis). Jadara University, Irbid, Jordan.
- Allam, S. M. (2007). *Measurement and educational evaluation in the teaching process*. Amman, Jordan: Dar Al-Masirah for Publishing and Distribution.
- Al-Najjar, N. (2010). *Measurement and evaluation: an applied perspective with SPSS software*. Amman, Jordan: Dar Al-Hamed for Publishing and Distribution.
- Al-Omari, R. (2015). *Extent of knowledge of science teachers of the basic stage in the directorates of education in the Northern Mazar brigade of the principles of building achievement tests*. (Unpublished doctoral thesis). Omdurman University, Sudan.

Al-Shayab, S. (2003). *The effectiveness of home economics teachers in building achievement tests in the Ministry of Education schools in Jordan according to good test criteria.* (Unpublished Master Thesis). Amman Arab University for Graduate Studies, Amman, Jordan.

Al-Shayab, S. et al. (2001). *Teacher guide in building achievement tests.* Examinations and Tests Department, Department of Educational Measurement and Evaluation, Ministry of Education, Jordan.

Al-Zahrani, S. (2001). *The degree of knowledge of teachers and supervisors at Qurayyat governorate to the principles of building achievement tests and practice them from their point of view.* (Unpublished Master Thesis). Yarmouk University, Irbid, Jordan.

Arabiyat, N. (2005). *The level of knowledge of physical education teachers in Balqa governorate of the concepts and terminology of measurement and evaluation and their estimation of the level of their use and their training needs.* (Unpublished Master Thesis). Amman Arab University, Amman, Jordan.

Bloom, B., S., et al. (1973). *Handbook on formative and summative evaluation of student learning*, 3rd ed. MCgraw-Hill: New York.

Boston, C. (2003). *High school report cards.* Eric digits, Eric clearing house on Assessment & Evaluation, 91, 660 -334.

Bothrogd, R. R. A. S. A., & Others. (1992). *What do teachers know about measurement and how did they find outs*, paper presented council on measurement in Education (San Francisco, CA).

Daniel, L. G., & King, D. A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *The Journal of Educational Research*, 91(6), 331-344. <https://doi.org/10.1080/00220679809597563>

Gronlund, E. (2000). *Measurement and evaluation in teaching.* Seventh Edition, New York, Mcmillan CO.

Huseyin, C. (2011). *Teacher opinion concerning common exams administrated in elementary schools.* (ERIC NO 520685).

in Jordan in the field of measurement and evaluation based on their training needs. (Unpublished Doctoral Thesis). University of Jordan, Amman, Jordan.

Melhem, S. (2002). *Measurement and evaluation in education and psychology* (Second ed.). Amman, Jordan: Dar Al-Masirah for Publishing and Distribution.

Metwally, A., Al-Tahawi, K., Qastanian, G. and Al-Rabai, A. (2005). The effectiveness of a computer-designed training program for the development of social culture among teachers in service in the Sultanate of Oman. *Prince Abdullah bin Abdulaziz Conference on Opinions on Higher Education.* Jeddah: King AbdulAzziz University.

Ministry of Education (2017). *Specialized content for vocational education course*, Amman: Jordan.

Odeh, A. (2010). *Measurement and evaluation in the teaching process* (fourth ed.). Irbid, Jordan: Dar Al Amal.

Otom, M. (2014). Evaluation of achievement tests prepared by science teachers in Jerash governorate: an analytical study of students' results for the second semester 2012-2013 in science. *Journal for Research and Studies*, 15(2), 361-375.

Saada, L. (2012). *A proposed program to develop the competencies of physical education teachers*

Simsk, A. (2016). A comparative analysis of common mistakes in achievements prepared by schoolteachers and corporate trainers. *European journal of science and mathematics education*, 4(4) 477-489.

Sinan, E. M. (2006). *The degree of knowledge of the practice of mathematics teachers in the intermediate stage of the skills of building achievement tests in Makkah Al-Mukarma*. (Unpublished Master Thesis). Umm Al-Qura University, Makkah, Saudi Arabia.

Copyright Disclaimer

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).