Dirasat, Educational Sciences, Volume **38**, No. **1**, 2011

Secondary School Students' Perceptions of Essay and Multiple-Choice Type Exams

Yahya Hayati Baker Nassar, Khalil Qaraeen and Abdullah Abu Naba'h *

ABSTRACT

This study aimed at investigating Jordanian secondary school students' perceptions in Amman toward essay and multiple-choice exams. The sample composed of 403 students (195 males and 208 females). The Inventory of Secondary Schools Perceptions of Essay and Multiple-choice Type Exams (ISSP) was constructed to measure students' perceptions. The results revealed that students preferred multiple-choice exam more than essay exam. Furthermore, students believed that multiple-choice exam is less difficult, less complex, clearer, more interesting, trickier, and fairer than essay exam. In addition students' believed that multiple-choice exam is better in terms the expectancy to success, less anxious, and better in terms feeling at ease comparing with essay exam. However, students' believed that both types are valuable. Moreover, the findings showed that higher achiever students prefer essay exam more than moderate and low achievers. Finally, students' perceptions toward multiple-choice exam did not differ significantly according to their gender and/or achievement level.

Keywords: Secondary School, Students' Perceptions, Evaluation, Essay Exam, Multiple-Choice Exam.

INTRODUCTION

The evaluation process can be considered as very important factor in education and psychology (Tozoglu, Tozoglu, Gurses and Dogar, 2004). It reflects the degree to which the objectives of curricula and other programs have been achieved (Airasian, 1997). The evaluation and assessment are used sometimes interchangeably and both of them refer to a related series of measures used to determine a complex attribute of an individual or group of individuals (Oosterhof, 2001). Students' perceptions regarding the different elements of the evaluation process can be considered very essential factor in terms of getting valid and reliable evaluation (Oosterhof, 2001). The valid evaluation of students' achievement, which reflects the outcomes of teaching and learning processes, can not be achieved without taking into account students' needs, students' perspectives, and their attitudes toward the different procedures or methods of evaluation (Nassar, 2004). Moreover, students' perceptions toward the different types of evaluation should be taken into account even if we want to train them to deal with the different types of exam.

Essay and multiple-choice type exams are the most popular item formats used in educational testing (Oosterhof, 2001, Tozoglu et al., 2004). Each of which required from the examinee to have a specific skills in order to get high scores (Nassar, 2006). Some students might be more able to perform on one of them better than the other one. So students' perceptions or attitudes toward the type of exam could differ according to their expectations about their performance on that exam. Moreover, although students sometimes report that multiple-choice exams are easier than essay exams, they in some situations could consider multiple-choice format more difficult than essay exam (Holtzman, 2008).

Usually, When the teachers in the schools construct achievement tests to evaluate their students, they take into account different factors such as the nature of the material, the available time to prepare the exam, the time which is needed to administer it, and the time is needed to the process of scoring. Depending on the previous factors and the experience of the teacher, his skills, and his believes, The teacher selects the type of the exam (Audeh, 1987). Moreover, the process of constructing a high quality test demands an understanding of the objectives and contents being assessed, taking into

* Department of Educational Psychology, Faculty of Educational Sciences, The Hashemite University, Alzarka, Jordan. Received on 19/3/2008 and Accepted for Publication on 31/12/2008.

account examinees' learning and thinking styles, examinees' reading level, and the degree of difficulty of the test (Hopkins, 1998; Huff, 1998). In addition to the mentioned factors, many studies indicated that students' attitudes and perceptions toward the different types of exams should be taken into account on the process of their evaluation (D'ydewalle, Swerts and De Cort, 1983; Gay, 1980; Hakstian, 1971; Nevo, 1985; Oosterhof, 2001; Tozoglu et al., 2004 and Zeidner; 1987). Since essay and multiple-choice exams are the most popular formats used to evaluate students, it is essential to investigate their perceptions toward both of them.

Essay exams include items or questions start with specific terms such as what, how, why, Justify, explain, and so on (Aldoghan, 1995). Essay exam required from the examinee to recall and to construct the answers, through their special way, for this reason this type of exams called in the literature constructed-response tests (Popham, 1981). Moreover there are advantages and disadvantages for essay type exam. For example the process of constructing its items does not take too much time or efforts comparing with the objective exams (Tuckman, 1993), it is essay to cover the higher thinking levels via this type of exam and an essay item could be considered as the best way to assess a student's ability to communicate ideas in writing (Tozoglu et al., 2004). On the other hand, it is difficult to cover the objectives of the course through this type of exam, the scoring process could be subjective and difficult, it is not easy to computerize such type of exam especially in terms of the scoring process, in addition there is a bluffing problem (Audeh, 1987). In general students do not prefer such type of exam for different reasons: they should construct the answer by themselves, it is need much time and efforts during the study, comparing with objective type exam, and they might fell that this type of exam is not fair in terms of their grades or scores (Zeidner, 1987).

Multiple-Choice tests also have advantages and disadvantages, so they easy to score, can be described objective in terms of their scoring, easy to computerize specially in term of scoring process, and more importantly it is easy to cover the objectives of the course through this type of exams (Ben-Simon, Budescu and Nevo, 1997). However, there are many problems associated with this type of exam such as the guessing problem, the difficulty of building items measure the advanced levels of thinking according to bloom's taxonomy, and the process of preparing its items is not easy and needs time and efforts (Oosterhof, 2001). However, some researchers argue that this type of exams can measure effectively the advanced mental processes (Mehrens and lehman, 1984). The literature indicates that students prefer this type of exams comparing with essay type exam (Zeidner, 1987; Choi, 1998; Birenbaum and Feldman, 1998) the explanation of these results might correlated with students believes about the process of preparing for this type of exam, so they perhaps think that with objective exams it is easier to study the material, and/or they might believe that this type of exam is more fair than essay type exam. The process of reviewing the literature about students' perceptions' of essay and multiple type exams indicates that few studies had been conducted to investigate those perceptions.

Zeidner (1987) tried to examine secondary school students' perceptions of essay and multiple choice type exams, the sample of his study consisted of 174 male and female students. Two instruments were applied in this study. The first one represented students' perceptions in terms of ten dimensions: difficulty, complexity, clarity, interest, trickiness, fairness, value, success expectancy, degree of anxiety evoked, and feeling at ease. The second instrument was used as criterion instrument, to examine the validity of the major instrument, this instrument included seven dimensions. In the second tool the students were asked to respond whether they prefer essay or multiple choice type exams. The results of this study revealed that secondary school students preferred multiple-choice exam more than essay exam.

In another study Aldoghan (1995) investigated King Saud University male students' attitudes toward essay and objective exams. Moreover, he attempted to examine whether students' attitudes' differ significantly according to their majors, academic level, and GPA. Research results showed that the students' attitudes toward objective tests were significantly more positive than their attitudes toward essay tests regardless of their majors, academic level, and GPA. So students' attitudes do not differ significantly according to the mentioned variables. In this context it should said that while students attitudes focus on their feelings about the object, their perceptions gave more important to how they perceive mentally that object.

In their study Tozoglu et al. (2004), examine students' perceptions toward essay versus multiple-choice exams. Fifty students from a science education department participated in this study. Overall student rating data revealed that students showed significantly ($p<0.001$)

more favorable attitudes toward multiple-choice test format compared to essay type formats in terms of the most critical dimensions assessed: difficulty, complexity, clarity, interest, trickiness, fairness, value, success expectancy, degree of anxiety evoked, and feeling at ease. Tozoglu et al. suggested in the findings of their study that students' perceptions should be taken into consideration while planning and constructing classroom testing.

In addition, according to the assumption that Students' perceptions about evaluation methods also play a significant role, Struven, Dochy, and Janssen (2005) examine students' perceptions regarding evaluation and assessment methods. Research findings reveal that students' perceptions about assessment significantly influence their approaches to learning and studying. on the other hand, students' approaches to study influence the ways in which they perceive evaluation and assessment. Moreover, Findings suggest that students hold strong views about different assessment and evaluation formats. In this respect, students favor multiple-choic***e*** format exams to essay type. However, when compared with more innovative assessment methods, students call the 'fairness' of these well-known evaluation modes into question.

Furthermore, Nassar (2006) aimed at exploring the effect of training the Hashemite University students on skills related to their achievement on true-false, multiple choice, completion, and essay items. In general he found that the training was useful in developing students' skills in dealing with the mentioned items. In his discussion of these results Nassar observed that the biggest difference was obtained in students' achievement means on essay items. He explained this result due to students' perceptions of essay items before and after the training. That is students might have negative attitudes toward this type of items because they do not have sufficient qualifications of dealing with such items.

Finally, Fornham, Christopher, Garwood, and Martin (2008) selected more than 400 students from four universities in America and Britain completed measures of learning style preference, general knowledge (as a proxy for intelligence), and preference for examination method. Learning style was consistently associated with preferences: surface learners preferred multiple choice and group work options, and viewed essay-type and dissertation options less favorably. Deep learners, on the other hand, favored essay-type and oral exams as well as final dissertations. Males favored oral exams and females coursework assessment. Extraverts preferred multiple choice, oral, and group work assessment, while openness was positively associated with essays and oral exams but negatively associated with multiple choice and group work. Regression analysis showed that personality, learning style, general knowledge, and demographic factors accounted for 5-10% of the variance in preferred examination technique.

On the light of the mentioned studies, educators should pay attention to students' perceptions because of their possible impacts on learning, the most important outcome of instruction (Tozoglu et al., 2004). Students' perceptions about a different educational situations can tell the teacher a grate deal about the impact of that situation on the learning process (Tozoglu et al., 2004). In addition, students' attitudes are associated with how they perceive or recognize the different situations in which they find themselves (Smith, 1982, Dusic, 1998). Although students' perceptions about the different types of exams can effect their performance on these exams (Nassar, 2004), there are just few studies ,specially in the secondary school level, have been conducted to investigate secondary school students perceptions about the different types of exams specially essay and multiple-choice type exams. Those studies focused on one or two dimensions such as test anxiety or general attitudes toward a test format (Aldoghan, 1995; Tozoglu et al., 2004). As a result, more attention should be given to examinees' perspectives because they play very essential factor in explaining students' performance on the test or on its components (Tozoglu et al., 2004). Zeidner (1987) indicated that students' perceptions' toward test format can be defined in terms of different components such as how they perceived interestingness, usefulness, trickiness, fairness and other elements.

Statement of the problem

The lack of studies in the area of students' perceptions toward essay and multiple-choice type exams could be applied to Jordanian secondary schools, so this study will try basically to examine secondary school students' perceptions in Amman toward essay and multiple-choice type exams and to investigate whether students' perceptions differ according to their gender and achievement level.

Research Questions:

This study will try to answer the following questions:

1- How secondary school students in Amman perceive the essay and multiple choice type exams?
2- Do secondary school students' perceptions of essay and multiple type exams differ according to their gender?
3- Do secondary school students' perceptions of essay and multiple type exams differ according to their achievement level?
4- Do secondary school students' perceptions of essay and multiple type exams differ according to the interaction between the variables gender and achievement level?

From the previous questions it could be observed that the current study differ from Zeidner's study and Tozoglu's study in its different cultural context and in its variables.

Table 1: The Number of Students in the Sample and Their Percentages According to the Levels of the Variables Gender and Achievement Levels

			Students Achievement Level			
			HIGH	MODERATE	LOW	Total
Students' Gender	male	Count	44	121	30	195
		% within Students' Gender	22.6%	62.1%	15.4%	100.0%
		% within Students Achievement Level	41.5%	49.6%	56.6%	48.4%
		% of Total	10.9%	30.0%	7.4%	48.4%
	female	Count	62	123	23	208
		% within Students' Gender	29.8%	59.1%	11.1%	100.0%
		% within Students Achievement Level	58.5%	50.4%	43.4%	51.6%
		% of Total	15.4%	30.5%	5.7%	51.6%
Total		Count	106	244	53	403
		% within Students' Gender	26.3%	60.5%	13.2%	100.0%
		% within Students Achievement Level	100.0%	100.0%	100.0%	100.0%
		% of Total	26.3%	60.5%	13.2%	100.0%

Limitations of the Study

The population of this study was males and female students' of the first grade in the secondary schools in Amman in the second semester of the academic year 2005-2006. Also the results of this study should be interpreted in the context of its definitions of the variables, its instrument, and the research method used to answer the research questions (survey method).

The Operational Definitions of the Variables:

Students' perceptions of essay and multiple-choice type exams: students scores on the instrument of secondary school students' perceptions of essay and multiple-choice type exams**.**

Students' achievement levels: Three judges of college of educational sciences at the Hashemite University were asked to classify students according to their average scores in the first semester of the academic year 2005- 2006 into three levels. As a result of that procedure students were classified as following:

High level achievers: Those students who their average ranged from 80-100.

Moderate level achievers: Those students who their average was more than or equal 65 and less than 80.

Low level achievers: Those students who their average was less than 65.

Methodology

Method of the study: in order to answer the questions of this study**,** the quantitative method, namely survey research, was used.

The population of the study: the population of the current study is males and females students in the secondary schools in the first and in the second Amman educational directorates (the first and the second educational districts) in the city of Amman.

Sample of the Study: the sample of this study consisted of 403 students (195 males and 208 females) of

the first grade in the secondary schools in the first and in the second Amman educational directorates. The multi-stage cluster random sampling procedure was used to select the sample. The number of students in the sample and their percentages according to the levels of the variables gender and achievement levels is shown in table (1).

In order to explain the numbers and the percentages in table (1), it might be useful to present an example. The number of male students who classified as high achievers was 44. This count of male students represents 22.6% from all of male students who participated in the study (195 students), 41.5% from the students who were classified as high achievers (106 students), and 10.9% from the whole sample (403 students). The rest counts and percentages in that table should be interpreted in the same manner.

The Instrument: In order to measure the perceptions' of the sample, which represented the secondary schools students in Amman, the Inventory of Secondary Schools Perceptions of Essay and Multiple-choice Type Exams (ISSP) was developed. This instrument was consisted from three sections, the first section includes the instructions, the second section includes two pieces of information: students' gender and their average score in the end of the first semester from the academic year 2005-2006, and the third section includes the items. The items of ISSP were prepared in Arabic by the researcher; the ideas of these items were taken from Zeidner's ten dimensions: difficulty, complexity, clarity, interest, trickiness, fairness, value (The importance), success expectancy, degree of anxiety evoked, and feeling at ease (Zeidner, 1987). Each one of those was measured via three items on each type of exams (essay and multiple-choice), so that the first three items of the items measuring students' perceptions toward essay type exam represent the first dimension (difficulty), and the next three items represent the second dimension (complexity), of the mentioned dimensions and so on. The same procedure was utilized with the items which constructed to measure students' perceptions' toward multiple-choice items. The instrument was constructed as a Likert-type rating scale on a five-point rang. ISSP composed of 60 Likert type items, 30 items to measure students' perceptions toward essay items and 30 items to measure their perceptions toward multiple choice items. The instrument included positive and negative items the positive items (40 items) were scored as following: Strongly Agree=5, Agree=4, Nuteron=3, Disagree=2, strongly Disagree=1. However the negative items (20 items) were scored in a reverse manner. Higher scores on the items indicate more favorable dispositions toward the test format under consideration. Based on the mentioned scoring method, a specific cut points to interpret students' total and sub-scores either those scores related to their perceptions toward essay exam or multiple-choice exam were prepared. So since students' perceptions toward essay exam were measured through 30 items and since the score of each of which was ranged from 1-5, the scale of the total scores of students' perceptions toward essay exam ranged from 30-150. The same scale was used to interpret students' perceptions toward Multiple-choice exam. The score 90, which is the middle point of that scale, was used to classify students' total scores into five categories. So total scores around 90 were considered as moderate while total scores that deviated more than or less than 15 units from the score 90 were considered either high or low. According to the mentioned criterion, the total scores in that scale were interpreted as following:

60 or less-very low, above 60 to 75-low, above 75 to 105-moderate, above 105 to 120-High, above 120 to 150-very high. Concerning the scale of students' scores on each dimension, it ranged from 3-15, the score 9 represents the middle point of that scale. The scores of that scale were interpreted as following: 6 or less-very low, above 6 to 7.5-low, above 7.5 to 10.5-moderate, above 10.5 to 12-High, above 12 to 15-very high.

The Arabic language was used in developing ISSP in order to be understandable from Jordanian students. Sample items from the instrument are as follows: "I believe that essay exam is fair", "I feel at ease with multiple-choice exam", "Multiple-choice exam increases my feelings of anxiety", "I believe that essay exam is difficult".

The Validity and Reliability of ISSP: For the purpose of examining the validity of the instrument (face validity evidences) it was presented to five experts in the fields of educational research and evaluation and educational measurement. They were asked to check whether the statements in the instrument are clear and linked appropriately with the dimensions they were classified to them in advance. The judges suggested few corrections such as adding one or two words to three items and rephrasing two items.

Regarding the reliability of the instrument two

procedures were used: Test-Retest procedure (to estimate stability coefficient) and internal consistency procedure (to estimate the consistency across the items). In order to estimate the reliability of the instrument, to analyze its items, and to examine the clearness of its statements and instructions, a pilot study had been conducted. 42 students in one class of eleventh grade (The first grade in the secondary stage) were utilized in that study. Those students did not participate in the final study. During the implication of the pilot study many students asked about the difference between the terms essay exam and multiple-choice exam. So it was important to make sure that students who will participate in the final study recognize the difference between those two types of exams before collecting the data. Except the final point the instructions were clear and all of the 60 items functioning in appropriate manner. Stability coefficients for the instrument in each case were 0.88 and 0.92 for essay and multiple-choice type of classroom test formats respectively. In addition, the values of alpha (the internal consistency coefficient) were 0.90 for essay exam and 0.97 for multiple-choice type exam. The previous values can be considered reasonably satisfactory to achieve the goals of the current study.

Collecting Data Procedure

In the beginning ten schools (5 form males' secondary schools and 5 from females' secondary schools) form the first and the second Amman educational directorates were selected randomly. From each selected school one class, from the eleventh grade classes in that school, was chosen randomly and all of the students who registered in that class were considered as a part of the sample of the study. The number of all of the students who were in those classes was 415. The instrument was administered to all of them, however, 12 copies of the inventory were excluded 4 of them because they left most of the items without responses and 8 of them because they were absent during collecting the data. As a result the final sample size for the present study was 403 students (195 males and 208 females). Students who participated in the sample were informed about the difference between essay and multiple-choice test formats before responding to the instrument. Moreover, as Tozoglu et al. (2004) did in their study, the students were told that the results of this study may encourage teachers to improve their classroom testing and they might give more attention in the future to students' perceptions' toward the different types of exams. Students respond to the instrument anonymously, and they need in average 20-25 minutes to accomplish the task. The process of collecting the data most of the time was done by the researcher himself or by some teachers who are working in the selected schools but with a clear instructions and a direct supervision from the researcher.

Data Analysis: The SPSS statistical package was used to analyze the data of the current study. In order to answer the first question of the study, the frequencies and the percentages of the secondary school students according to their perceptions toward essay and multiple-choice type exams were obtained. Also the first question was answered via the means and the standard deviations of the total scores of students' perceptions toward essay and multiple-choice type exams. And to examine whether there is a significant difference between the means of the total scores of students' perceptions toward essay exam versus Multiple-choice exam, the paired samples t- test was carried out. The same statistical procedure was utilized to examine if there are a significant differences between the means of the scores on each dimension of students' perceptions' of essay type exam versus multiple-choice type exams. And to answer the other questions of the current study the descriptive and the inferential statistical results of the procedure two-way ANOVA were employed. Moreover, these results were indicated to the effect size estimates. Finally, the research questions from 2 to 4 were answered via using (0.05) as level of significance.

RESULTS AND DISCUSSION

In order to describe secondary school students' perceptions in Amman toward essay and multiple –choice type exams, tables 2 and 3 revealed the frequencies and the percentages of their perceptions toward those types of exams.

As shown in table 2 (16.4%) from secondary school students' in Amman their perceptions toward essay exam can be described very low, were as 36% their perceptions were considered low. The last result indicated that students' perspectives might overestimate the negative points of that type of exam and underestimate it positive points. Table 2 also revealed that (18.1%) from the students' their perceptions toward that type of exam can be described moderate. That mean those students perceive the positive and the negative points of that type of exams in an equal manner. The important result of that

table 2 was that (23.6%) from the students had overestimated perceptions to the positive points of essay exam and they had underestimated perceptions to its negative points. This result might reflect that Jordanian secondary school students realize the value of this type of exam although most of them preferred, as revealed in table 3, to be examined via multiple- choice type of exam.

Table 2: The Frequencies and the Percentages of Secondary School Students' Perceptions in Amman Toward Essay Exam

		Frequency	Percent	Cumulative Percent
Valid	Very Low	66	16.4	16.4
	Low	145	36.0	52.4
	Moderate	73	18.1	70.5
	High	24	6.0	76.4
	Very High	95	23.6	100.0
	Total	403	100.0	

Table 3: The Frequencies and the Percentages of Secondary School Students' Perceptions in Amman Toward Multiple-Choice Exam

		Frequency	Percent	Cumulative Percent
Valid	High	66	16.4	16.4
	Very High	337	83.6	100.0
	Total	403	100.0	

Table 4: The Means and the Standard Deviations of the Total Scores of Secondary School Students' Perceptions toward Essay and multiple-choice exams

	Mean	N	Std. Deviation
total-essay	85.4553	403	35.70013
total-multiple	129.2231	403	8.97888

Table 3 revealed that (16.4%) of the secondary school students their perceptions toward the multiple-choice exam can be described as high, while (83.6%) their perceptions toward that type of exams can be described as very high. These results indicated that students had overestimated perceptions to the positive points of multiple-choice exam. However, the interesting result of table 3 is that no one of the sample his or her perceptions toward multiple-choice exam can be described as Low or even moderate. The last result might indicate that most of the secondary school students in Amman prefer such type of exam. Table 3 results could be interpreted by taking into the account that the psychological situation of Jordanian students, so they might feel less anxious with multiple-choice exam comparing with essay exams.

For the purpose of more explanation of secondary school students' perceptions in Amman toward essay exam and multiple- choice exams, some descriptive statistical results were obtained and table 4 shows these results.

As revealed in table 4 the mean of the total scores of students' perceptions toward multiple-choice exam is

greater than the mean of the total scores of their perceptions toward essay exam. Moreover, table 4 shows that students' perceptions toward essay and toward the multiple-choice type exams can be described as moderate and very high respectively. The obtained result regarding students' perceptions toward multiple-choice exam could be explained in the light of the literature (Zeidner, 1987; Choi, 1998; Birenbaum and Feldman, 1998) which indicates that this type of exam has different positive points, at least from students' point of views. However, students' have reasonable perceptions toward essay exam, which means that they were able not only to recognize the negative points of that type, but also they were able to identify the positive points associated with it.

Table 5: The Means and the Standard Deviations of the Total Scores of the Ten Dimensions of Students' Perceptions toward Essay and Multiple-Choice Exams

		Mean	N	Std. Deviation
Pair 1	Essay- Dificulty	8.1166	403	3.81002
	Multiple- Dificulty	13.1836	403	1.19744
Pair 2	Essay-Complexity	8.4194	403	3.93167
	Multiple-Complexity	13.0620	403	1.36539
Pair 3	Essay-Clarity	8.3747	403	3.97893
	Multiple-Clarity	13.2382	403	1.09628
Pair 4	Essay-Interest	8.3697	403	3.96813
	Multiple-Interest	13.1117	403	1.01968
Pair 5	Essay-Trickiness	8.3151	403	4.04298
	Multiple-Trickiness	13.0993	403	1.31394
Pair 6	Essay-Fairness	8.1737	403	3.96372
	Mutiple-Fairness	13.1191	403	1.26204
Pair 7	Essay-Value	10.7035	403	1.97485
	Multiple-Value	10.8757	403	3.88004
Pair 8	Essay-Success Expectancy	8.2432	403	3.95754
	Multiple-Success Expectancy	13.1687	403	1.17024
Pair 9	Essay-Anxiety	8.3648	403	4.03202
	Multiple-Anxiety	13.2109	403	1.21462
Pair 10	Eassy- Feeling at Ease	8.3747	403	3.98767
	Multiple- Feeling at Ease	13.1538	403	1.15735

In order to examine whether there is a significant difference between the means of the total scores of secondary school students' perceptions in Amman toward essay exam and multiple- choice exam, the dependent (t) test was carried out.

The results of dependent (t) test indicate that the difference between the means of the total scores of secondary school students' perceptions in Amman toward essay and multiple-choice exams is significant ($p<0.001$). This result and table 4 results revealed that the mean of students' perceptions toward multiple-choice exam is significantly greater than the mean of the total scores of their perceptions toward essay exam. The previous results are at the same direction with the results were mentioned by Zeidner (1987),Aldoghan (1995), and Tozoglu et al., (2004). These results might sustain that secondary school students in Amman feel more comfortable with multiple-choice of exam than essay exam. Students might believe that multiple-choice exam is easer than essay exam either in terms of recalling the information or responding to the item, since it does not need constructing the response as the case in essay items. The previous results and other reasons might make students feel less anxious with multiple-choice exam as mentioned before.

Table 6: The Dependent (t) Test Results for the Differences between the Means of the Total Scores on Each Dimension of Students' Perceptions toward Essay and Multiple-Choice Exams

		Paired Differences				
		Mean	Std. Deviation	t	df	Sig. (2-tailed)
Pair 1	Essay- Dificulty - Multiple- Dificulty	-5.067	4.22595	-24.070	402	.000
Pair 2	Essay-Complexity - Multiple-Complexity	-4.643	4.38976	-21.231	402	.000
Pair 3	Essay-Clarity - Multiple-Clarity	-4.864	4.24015	-23.026	402	.000
Pair 4	Essay-Interest - Multiple-Interest	-4.742	4.33290	-21.970	402	.000
Pair 5	Essay-Trickiness - Multiple-Trickiness	-4.784	4.39929	-21.831	402	.000
Pair 6	Essay-Fairness - Mutiple-Fairness	-4.945	4.34025	-22.874	402	.000
Pair 7	Essay-Value - Multiple-Value	-.17221	3.01027	-1.148	402	.251
Pair 8	Essay-Success Expectancy - Multiple-Success Expectancy	-4.926	4.18709	-23.615	402	.000
Pair 9	Essay-Anxiety - Multiple-Anxiety	-4.846	4.27257	-22.770	402	.000
Pair 10	Eassy- Feeling at Ease - Multiple- Feeling at Ease	-4.779	4.22070	-22.731	402	.000

Concerning the interpretation of the total scores of the ten dimensions of secondary school students' perceptions in Amman toward essay and multiple-choice exams, first the means and standard deviations of students' perceptions toward essay and multiple-choice exams across the ten dimensions were attained, and table 5 shoes these results.

Table 5 reveals that in most of the cases the mean of students' perceptions toward multiple-choice exam was grater than their mean toward essay exam. In order to make correct understand to the results of table 5, It should be observed that students' high scores on the dimensions difficulty, complexity, and anxiety do not mean that they believe that the specific type of exam (e.g. multiple-choice exam) is more difficult, complex, or anxious. Vice versa such scores indicate that students think that the specific type of exam is less difficult, complex, or anxious. The positive items of these dimensions reflect that meaning, while the negative items such as the following items "Multiple-choice exam increases my feelings of anxiety", "I believe that essay exam is difficult", reflect the opposite meaning. So, According to the results of table 5 secondary school students in Amman believe that multiple-choice exam is less difficult, less complex, clearer, more interesting, trickier, and fairer than essay exam. In addition students' believed that multiple-choice exam is better in terms the expectancy to success, less anxious, and better in terms feeling at ease comparing with essay exam. Tozoglu et al. (2004) indicated to similar results in their study except that their sample believed that multiple-choice exam is less tricky. The difference in the results of the current study and the results of Tozoglu et al., study regarding trickiness dimension might existed because the sample of this study represented the secondary school while the sample of Tozoglu et al., study represented university students. University students might be more skillful in terms of dealing with multiple-choice exam than secondary school students. In addition, the results of table 5 have different applications some of which related to the students. The students should be trained to deal with such items (Nassar, 2006). And they should realize that the

highest thinking levels could be measured via this type of exam, so they should take it seriously through spending more time and efforts to prepare to such exam. Also the teachers should be trained about the different skills related to the process of constructing such exam. And teachers should have a bank of multiple-choice items to be able to prepare different versions for the same exam.

Table 7: Descriptive Statistical Results Concerning Students' Perceptions toward Essay Exam According to their Gender and Achievement Level

Dependent Variable	Gender	Achievement Leve	Mean	Std. Deviation	N
Students' Total Scores of Their Perceptions Toward Essay Exam	Male	High	115.67045	30.602657	44
		Moderate	76.19835	34.756203	121
		Low	69.20000	19.735143	30
		Total	84.02821	36.226172	195
	Female	High	112.90323	27.051687	62
		Moderate	76.76423	33.786961	123
		Low	70.04348	23.340325	23
		Total	86.79327	35.234625	208
	Total	High	114.05189	28.470028	106
		Moderate	76.48361	34.201589	244
		Low	69.56604	21.162839	53
		Total	85.45533	35.700129	403

Table 8: Two Way ANOVA Results for the Effect of Secondary School Students' Gender, Achievement Level, and the Interaction between the Two Variables on Students' Perceptions toward Essay Exam

Dependent Variable	Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Students' Total Scores of Their Perceptions Toward Essay Exam	Gender	13.962	1	13.962	.014	.905	.000
	Achievlevl	118271.295	2	59135.647	59.826	.000	.232
	Gender * Achievlevl	219.237	2	109.618	.111	.895	.001
	Error	392419.050	397	988.461			
	Total	3455302.250	403				

Moreover, to examine whether there are a significant differences between the means of the total scores on each dimension of secondary school students' perceptions in Amman toward essay exam and multiple- choice exam, the dependent (t) test was employed ten times. However, the results were examined at (.005) instead of (.05) to reduce the probability of the inflation of type one error. The last procedure is called the Bonferroni adjustment, according to that procedure the expected level of significance should be divided on the number of times the same statistical procedure will be used in the process of analyzing the data (Howell, 1995). So in the current analysis (0.05) was divided over (11) and the result was (.0045). Table 6 reveals the results of dependent (t) test.

Table 6 shows that all of the differences between the means of the total scores on each dimension of students' perceptions toward essay and multiple-choice exams were significant at 0.0045 as level of significance except on the dimension value or the importance. So the secondary school students in Amman do not believe that the essay exam is less valuable than multiple-choice exam. That result could be explained through students' believes about essay exam. So, although students believed that multiple-choice exam was fairer especially in terms decisions concerning passing and failing (Tozoglu et al., 2004). They might think that essay exam is like multiple-choice exam in terms of its ability to reflect their knowledge. Another explanation is that essay

items allow students construct their own responses, develop a rational to support their thinking and position taken, require students to select and organize the content for discussion, consider multiple perspectives and present their ideas logically (Oermann, 1999).

In terms of examining whether secondary school students' perceptions toward essay and multiple-choice exams differ significantly at .05 level of significance according to their gender, achievement level, and/or the interaction between gender and achievement level, the statistical procedure Two-Way ANOVA was used.

Concerning Two-Way ANOVA results with students' perceptions toward essay exam as dependent variable was carried out. Table 7 shows the descriptive statistical results associated with that analysis.

Table 7 reveals that males and females students' perceptions toward essay exam were affected by their achievement level. So, the higher the achievements level, the greater the mean of the total scores of students' perceptions toward essay exam. However, the difference between students' means according to their gender was less than the revealed differences within the variable students' achievement level. The highest mean of students' perceptions toward essay exam (115.67) was to male students' who were classified as high achievers. While, the lowest mean of students' perceptions toward essay exam (69.20) was to male students' who were classified as low achievers. Moreover, In general female students' mean was grater than male students' mean in terms of their perceptions toward essay exam. The last result might indicates to that female students prefer such type of exam because they are more skilled than male students in terms of recalling , organizing , and expressing the meaning of a specific content.

In order to examine whether the revealed differences among the means in table 8 are significant at .05 level of significance, two-way ANOVA statistical procedure was used. Table 8 shows the obtained results.

Table 8 demonstrates that students' perceptions toward essay exam did not differ significantly ($p>0.05$) according to their gender or according to the interaction between students' gender and their achievement level. However, the results of table 8 revealed significant differences ($p<0.001$) among the means of total sores of students' perceptions toward essay exam. Moreover, according to Cohen (1977) the effect size magnitude should be considered as small at .01, moderate at .06, and large at .14. In the light of the previous criteria, table 8 indicates to very small effect size either to the variable gender or to the interaction between the variables students' gender and their achievement level regarding their effects on secondary school students' perceptions toward essay exam. However, table 8 shows that the effect size magnitude (.232) of the variable students' achievement level could be described as large. The last result indicates that the effect of students' achievement on their perceptions is not only statistically significant but also it is practically significance.

For the purpose of examining the differences among students' means of their total scores regarding their perceptions toward essay exam according to their achievement level, Tukey test (Multiple comparison procedure) was used. This procedure was selected because it is not extremely conservative (e.g. Schefee test) or extremely liberal (e.g. LSD test) (Howell, 1995).

The results of Tukey test indicate that the means of the total scores of high achievers of secondary school students regarding their perceptions toward essay exam were grater significantly ($p<0.001$) than the means of total scores of moderate and low achievers students regarding their perceptions toward essay exam. The previous results might indicate that the high achiever students may prefer such type of exam because it allows them to express their deep understanding of the content, knowledge, and ideas in writing (Tozoglu et al., 2004).

Concerning examining whether secondary school students' perceptions toward multiple-choice exams differ significantly at (0.05) level of significance according to their gender, achievement level, and/or the interaction between gender and achievement level, the statistical procedure Two-Way ANOVA was used. Table 9 illustrates the descriptive statistical results associated with that analysis.

Table 9 exposes that students' perceptions toward multiple-choice exam did not affected by their gender and/or their achievement level. All of the revealed means in table 9 can be described as high. This means that in general secondary school students in Amman have positive perceptions toward this type of exam. That result might explained via students' total scores in the ten dimensions so they believe that this type is less difficult, less complex, more clear, more interesting, more fair, more valuable, better in terms the expectancy to success, less anxious, and better in terms feeling at ease comparing with essay exam. The highest mean of students' perceptions toward Multiple-Choice exam

(129.97) was to female students' who were classified as moderate achievers. While, the lowest mean of students' perceptions toward multiple-choice exam (127.24) was to male students' who were classified as high achievers. It is very clear that the gab between the mentioned means is very tiny.

Table 9: Descriptive Statistical Results Concerning Students' Perceptions toward Multiple-Choice Exam According to their Gender and Achievement Level

Dependent Variable	Gender	Achievement Level	Mean	Std. Deviation	N
Students' Total Scores of Their Perceptions Toward Essay Exam	Male	High	127.2477	10.49397	44
		Moderate	129.5455	8.14760	121
		Low	127.7000	6.93890	30
		Total	128.7431	8.58749	195
	Female	High	129.4839	7.85262	62
		Moderate	129.9756	9.75559	123
		Low	128.5652	10.87460	23
		Total	129.6731	9.32913	208
	Total	High	128.5557	9.06352	106
		Moderate	129.7623	8.97830	244
		Low	128.0755	8.77902	53
		Total	129.2231	8.97888	403

Table 10: Two Way ANOVA Results for the Effect of Secondary School Students' Gender, Achievement Level, and the Interaction Between the two variables on students' Perceptions toward Multiple-choice Exam

Dependent Variable	Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Students' Total Scores of Their Perceptions Toward Multiple-Choice Exam	Gender	94.437	1	94.437	1.169	.280	.003
	Achievlevl	207.572	2	103.786	1.285	.278	.006
	Gender * Achievlevl	59.148	2	29.574	.366	.694	.002
	Error	32071.673	397	80.785			
	Total	32409.355	402				

In order to examine whether the revealed differences among the means in table 9 are significant at (0.05) level of significance, Two-Way ANOVA statistical procedure was used. Table 10 shows the obtained results.

Table 10 reveals that students' perceptions toward multiple-choice exam did not differ significantly ($p>0.05$) according to their gender, achievement level, or according to the interaction between students' gender and their achievement level. Moreover, table 10 shows very small effect size either to the variable gender, students' achievement level, or to the interaction between the variables students' gender and their achievement level (.003, .006 and .002 respectively) regarding their effects on secondary school students' perceptions toward multiple-choice exam.

CONCLUSION AND RECOMMENDATIONS

Students' perceptions of essay and multiple-choice exams can be considered one of the elements of that comprehensive process which called evaluation. As founded in this study there is a relationship between students' level of achievement and their perceptions at least toward essay exam. It seems that high achiever students prefer this type of exam more than lower achievers. The previous result might indicate that high achiever students are less concern about the type of exam than lower achievers students. However, the results of the

present study found that secondary school students' perceptions toward multiple-choice exam did not effected by their gender or their achievement level. It is very obvious either through the results of this study or the results of the other studies in literature that students in general have positive attitudes toward multiple-choice exam. These results might due to the advantages of this type of exam, such as its objectivity, it validity in terms of covering most of the content or the objectives of the course, and students' felling at ease with that type of exam. The interesting result was that although students prefer multiple-choice exam more than essay exam, they believed that they do not differ in terms of their value or importance. This result reflect that students aware about the different advantages of essay and multiple-choice type of exams.

Due to the results of the current study and to mentioned results in the literature, the following recommendations could be presented:

1- School teachers should take into account their students perceptions toward the different type of exams. Since students' scores on these exams are used to take decision about their level of achievement.
2- As Nassar (2006) mentioned students' should be trained to deal with the different types of exams. So students should have the sufficient skills to deal with these types.
3- In order to make the students ready to deal with the different types of exams, teachers should use these different types from early stages (From the elementary stage) rather than focusing only on one type.
4- To be able to take into account their students perceptions regarding the exams, teachers should have the necessary skills to be able to build different types of exams.
5- Since the results of this study revealed that students' perceptions of essay exam differ according to their achievement levels, teachers should take into account that factor when they prepare their exams.

REFERENCES

Airasan, P.W. 1997. *Classroom Assessment* (3rd ed.). New York: McGraw-Hill.

Aldoghan, A.A. 1995. King Saud university students' attitudes toward essay and objective tests. *Journal of King Saud University: Educational Sciences and Islamic Studies,* 7: 275-312.

Audeh, A. 1987. *Measurement and Evaluation In The Process of Teaching*. Irbid: Dar Alamal.

Ben-Simon, A., Budesco, D.V. and Nevo, B. 1997. A comparative study of measures of partial knowledge in multiple-choice tests. *Applied Psychological Measurement,* 21 (1): 65-88.

Birenbaum, M., Feldman, R.A. 1998. Relationships between learning patterns and attitudes toward two assessment formats. *Educational Research,* 40: 90-98.

Choi, N. 1998. The effects of test format and locus of control on test anxiety. *Journal of College Student Development,* 39: 616-620.

Cohen, J. 1977. *Statistical power analysis for the behavioral sciences* , New York: Academic Press.

Dusic, D.M. 1998. What social cognitive factors influence faculty members' use of computers for teaching? A literature review. *Journal of Research on Computing in Education,* 31: 123-137.

D'Ydewalle, G. Swerts, A. and De cort, E. 1983. Study time and test performance as a function of test expectations. *Contemporary Educational Psychology,* 8:55-67.

Furnharm, A., Christopher, A., Garwood, J. and Martin, N.G. 2008. Ability, demography, learning style, and personality trait correlates of student preference for assessment method. *Educational Psychology,* 28 (1): 15-27.

Gay, L. R. 1980. The comparative effects of multiple-choice versus short-answer tests on retention. *Journal of educational measurement,* 17: 45-50.

Hakstian, R. 1971. The effects of type of examination anticipated on test preparation and performance. *Journal of Educational Research,* 64: 319-324.

Holtzman, Mellisa. 20088. Demystifying application-based multiple-choice questions. *College Teaching,* 56 (2): 114-120.

Hopkins, K. 1998. *Educational and Educational Measurement and Evaluation*. Nedham Heights, MA: Allyn and Bacon.

Howell, D.C. 1995. *Fundamental statistics for the behavioral sciences* (3rd ed.). CA: Belmont.

Huff, J. A. 1998. *Assessment in science: An examination of multiple-choice and performance science assessments for a second grade classroom*. (ERIC Document Reproduction Service NO. ED 423 286).

Mehrens, W. A. and Lehman, I. J. 1984. *Measurement and Evaluation in Education and Psychology*. (3rd ed.). New York, Holt, Rinehart and Winston.

Nassar, Y. H. 2004. King Saud university students' perceptions regarding evaluation methods and exams as learning tool and their relationship with students' specializations and achievement levels. *The Educational Journal*, 18 (70): 11-64.

Nassar, Y. H. 2006. The effect of training the Hashemite University students on some skills of dealing with some item formats on their academic achievement. *The Educational Journal*, 20 (79): 69-119.

Nevo, B. 1985. Face validity revisited. *Journal of Educational Measurement*, 22: 287-293.

Oermann, M. 1999. Developing and Scoring Essay Test. *Nurse Educator*, 24 (2): 29-32.

Oosterhof, A. 2001. *Classroom applications of educational measurement*. Upper Saddle River, New Jersey, Columbus, Ohio: Merill Prentice Hall, Inc.

Popham, W. J. 1981. *Modern Educational Measurement*. New Jersey, Prientice Hall, Inc.

Struven, K., Dochy, F. and Janssen, S. 2005. Students' perceptions about evaluation and assessment in higher education: A review. *Assessment and Evaluation in Higher Education*, 30 (4): 325-341.

Tozoglu, D., Tozoglu, M. D., Gurses, A. and Dogar, C. 2004. The students' perceptions: Essay versus multiple-choice type exams. *Journal of Baltic Science Education*, 2 (6): 52-59.

Tuckman, B.W. 1993. The essay test: a look at the advantages and disadvantages". *Nassp Bulletin*, 77 (55): 20-26.

Zeidner, M. 1987. Essay versus multiple-choice type classroom exams: The students perspectives, *Journal of Educational Research*, 80 (6): 352-358.

إدراكات طلبة المدارس الثانوية حول الاختبار المقالي واختبار الاختيار من متعدد

يحيى نصار، وخليل قراعين، وعبد الله أبو نبعة*

ملخـص

هدفت هذه الدراسة لفحص إدراكات طلبة المدارس الثانوية في الأردن حول الاختبار المقالي والاختبار من نوع الاختيار من متعدد. تكونت عينة الدراسة من 403 طلاب (195 ذكور و208 إناث). وقد تم بناء قائمة إدراكات طلبة المدارس الثانوية حول الاختبار المقالي والاختبار الموضوعي لقياس إدراكاتهم. أشارت نتائج الدراسة إلى أن الطلبة يفضلون بصورة عامة الاختبار من نوع الاختيار من متعدد أكثر من الاختبار المقالي، ويعتقدون أن الاختبار من نوع الاختيار من متعدد هو اقل صعوبة وأقل تعقيداً وأكثر وضوحاً وأكثر تشويقاً وأكثر خداعاً وأكثر عدالةً وأفضل من حيث توقع النجاح وأقل إثارة للقلق وأفضل من حيث الشعور بالارتياح. بينما أشار الطلبة إلى أن كلا النوعين يعتبر ذا قيمة. وكذلك أشارت النتائج إلى أن الطلبة ذوي المستوى التحصيلي المرتفع هم الأكثر تفضيلاً للاختبار المقالي. وأخيراً لم تختلف إدراكات الطلبة حول اختبار الاختيار من متعدد بصورة دالة إحصائياً تبعاً لمتغيري الجنس والمستوى التحصيلي.

الكلمات الدالة: المدارس الثانوية، إدراكات الطلبة، التقويم، الاختبار المقالي، اختبار الاختيار من متعدد.

* كلية العلوم التربوية، الجامعة الهاشمية، الأردن. تاريخ استلام البحث 2008/3/19، وتاريخ قبوله 2008/12/31.