



ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΣΤΑ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

Διπλωματική Εργασία

STUDENTS' ATTITUDES TOWARDS ALTERNATIVE MULTIPLE CHOICE QUESTIONS

του

ΛΕΩΝΙΔΑ ΟΞΟΥΖΟΓΛΟΥ

Υποβλήθηκε ως απαιτούμενο για την απόκτηση του μεταπτυχιακού
διπλώματος ειδίκευσης στα Πληροφοριακά Συστήματα

Φεβρουάριος 2013

Στην Αμαλία . . .

Στην Ευτυχία . . .

Στην Ειρήνη . . .

Ευχαριστίες

Ευχαριστώ τον καθηγητή μου κ. Αναστάσιο Οικονομίδη για την πολύτιμη καθοδήγηση και την υποστήριξη στη συγγραφή αυτής της εργασίας.

Ευχαριστώ τον φίλο και συνάδελφο μου κ. Κώστα Σταυρίδη για τη βοήθεια και την υποστήριξη.

ABSTRACT

Multiple Choice Questions is a convenient and cost-effective way of assessing student ability. Issues concerning the assessment of partial knowledge and the use of guessing by examinees have led to the proposal of alternative methods of Multiple Choice testing. In the present study, students' attitudes towards six assessment methods were measured. Two of the methods are conventional and widely used while the rest are methods of alternative Multiple Choice Questions. The sample was students from both the secondary education and university and attitudes towards the methods were assessed with the use of questionnaires. The results showed a strong relation between perceived fairness and perceived accuracy in measuring a students' knowledge. Both of them are a decisive factor in students' preference in a method. Regarding preference, the conventional methods (essay type questions and conventional multiple choice questions) seem to be the students' choice over the alternative methods. A strong disapproval of multiple choice questions with negative marking was evident. Indications of the significance of certain factors (Approach to Learning, Educational Level, Preference in Courses, School Performance) in shaping assessment preference are also presented and discussed.

TABLE OF CONTENTS

1. Introduction	1
2. Review of Literature	2
2.1. Methods of testing with Multiple Choice Questions	2
2.1.1. Number of Right (NR) Scoring	2
2.1.2. Correction for Guessing methods	2
2.1.3. The Answer Until Correct method (AUC)	4
2.1.4. Elimination Testing (ET)	4
2.1.5. Subset Selection Testing (SST)	5
2.2. Students Assessment Preferences	7
3. Purpose of the Study	8
4. Method	9
4.1. Participants	9
4.2. Questionnaires	9
4.2.1. Approaches to Learning	9
4.2.2. Preferred Assessment Method	10
4.3. Procedure	10
5. Results	11
5.1. Overall Results	11
5.2. Secondary Student Results	27
6. Conclusion	30
7. References	33
APPENDIX	36
Questionnaire Used for Secondary Students	36
Questionnaire Used for University Students	40

LIST OF TABLES

Table 1: Cronbach's alpha of both scales of the R-LPQ-2F and R-SPQ-2F questionnaires.....	9
Table 2: Mean scores and SDs of preferred assessment methods	12
Table 3: Pairwise comparison of the mean values of Level of Preference using Repeated measures ANOVA	13
Table 4: Means of Level of Preference categorized by gender.....	14
Table 5: Means of Level of Preference categorized by grade	15
Table 6: Correlations between Level of Preference in different methods.....	16
Table 7: Correlations between Level of Preference in each method and learning approach scale scores	17
Table 8: Correlations between Level of Preference in each method and difference between learning approach scale scores	17
Table 9: Means of Level of Preference categorized by Learning Approach.....	18
Table 10: ANOVA of Level of Preference in Essay Type Questions with Learning Approach and Gender factors.....	19
Table 11: Mean values of Level of Preference in Essay Type Questions with the sample categorized by Learning Approach and Gender.....	19
Table 12: ANOVA of Level of Preference in AUC with Learning Approach and Gender factors .	20
Table 13: Mean values of Level of Preference in AUC with the sample categorized by Learning Approach and Gender	21
Table 14: ANOVA of Level of Preference in SST with Learning Approach and Gender factors..	22
Table 15: Mean values of Level of Preference in SST with the sample categorized by Learning Approach and Gender	22
Table 16 : ANOVA of Level of Preference in SST with Learning Approach and Grade factors....	23
Table 17: Mean values of Level of Preference in SST with the sample categorized by Learning Approach and Grade	24
Table 18: ANOVA of Level of Preference in NR with Learning Approach, Gender and Grade factors	25
Table 19: Mean values of Level of Preference in NR with the sample categorized by Learning Approach, Gender and Grade	25
Table 20: ANOVA of Level of Preference in ET with Learning Approach, Gender and Grade factors	26

Table 21: Mean values of Level of Preference in ET with the sample categorized by Learning Approach, Gender and Grade	26
Table 22: Correlations between Level of Preference in each method and school performance	27
Table 23: Means of Level of Preference with the sample categorized by Preference in Mathematics	28
Table 24: Means of Level of Preference with the sample categorized by Preference in Science Courses.....	29
Table 25: Means of Level of Preference with the sample categorized by Preference in Language Courses.....	29

LIST OF GRAPHS

Graph 1: Means of students attitudes towards the six methods of assessment.....	12
Graph 2: Means of Level of Preference for the two genders	14
Graph 3: Means of Level of Preference of Secondary and University students	15
Graph 4: Means of Level of Preference of students with a Deep Learning Approach and Surface Learning Approach	18
Graph 5: Boxplots of Level of Preference in Essay Type Questions with the sample categorized by Learning Approach and Gender	20
Graph 6: Boxplots of Level of Preference in AUC with the sample categorized by Learning Approach and Gender	21
Graph 7: Boxplots of Level of Preference in SST with the sample categorized by Learning Approach and Gender	23
Graph 8: Boxplots of Level of Preference in SST with the sample categorized by Learning Approach and Grade	24
Graph 9: Means of Level of Preference of secondary students categorized by preference in Mathematics	28
Graph 10: Means of Level of Preference of secondary students categorized by preference in Science Courses.....	29
Graph 11: Means of Level of Preference of secondary students categorized by preference in Language Courses.....	30

1. Introduction

The most common method of written assessment, especially when dealing with a large number of candidates is the Multiple Choice Questions method. MCQs are widely used in standardized summative assessment (SAT, ACT, GRE), in assessment of foreign language skills and also in education of any level (DiBattista & Kurzawa, 2011; Siegfried et al., 1996). Their use started in the USA in the first quarter of the 20th century and since then they are broadly used in all parts of the world.

There are many advantages in using MCQs in written assessment. MCQs are objective, easy to score (both by humans or computer-based applications) and their use offers significant cost-savings especially in large scale assessment (Shaftel & Shaftel, 2007; Williams, 2007). The sampling of content can be broader than in constructed response tests (Shaftel & Shaftel, 2007) and items can be stored in item-banks and reused (Haladyna & Downing, 1989). Individual test items can be subjected to statistical analysis and poorly performing items can be removed from the test (Shaftel & Shaftel, 2007). Moreover, the use of large number of items is possible, contributing to test reliability (Haladyna & Downing, 1989).

On the other hand MCQs are more difficult to construct (Haladyna & Downing, 1989). It is also possible that the examinee is exposed to acquiring false knowledge by remembering the distractors (Roediger & Marsh, 2005). Another point of criticism of MCQs is that despite the fact that this kind of testing has reliable results when assessing low level cognitive skills, when used for assessing higher levels, the reliability is questioned (Shaftel & Shaftel, 2007; Simkin & Kuechler, 2005). Criticism is also centered upon the fact that the score might be affected by the use of guessing (Akeroyd, 1982; Budescu & Bar – Hillel, 1993; Burton, 2001). Moreover, another point of discussion is the lack of assessment of partial knowledge (Coombs et al., 1956; DeFinneti, 1965; Bradbart et al., 2004). In other words, although an examinee cannot identify the correct response, he/she quite often can determine that some options are incorrect.

To address the problems of guessing and assessment of partial knowledge, many alternative methods of testing using MCQs have been

proposed. In the next session there is a brief presentation of these methods and an overview of the related literature.

2. Review of Literature

2.1. Methods of testing with Multiple Choice Questions

2.1.1. Number of Right (NR) Scoring

Number of Right (NR) scoring in testing with MCQs (also referred to as Number of Correct scoring or Number Scoring) is the scoring of a test just by counting the number of correct answers. If a wrong option is chosen, the score of the item is 0. NR scoring is the most common method of testing using MCQs. Criticism upon the NR method tends to center upon the fact that the score is affected by guessing and no credit is given for partial knowledge (Akeroyd, 1982; Bradbart et al., 2004). There is also discussion about the reliability of NR, especially in tests with a small number of items or items with a small number of options.(Burton, 2005)

2.1.2. Correction for Guessing methods

The oldest (Trow, 1923) and more common method for correction for guessing in testing using MCQs is that of Negative Marking (NM). In this method the wrong choice is given a negative score while the right choice a positive one. The scoring rule that is widely used in NM is to credit the wrong answer with $-\frac{1}{k-1}$ of the overall score of the item, where k is the number of choices of a MCQ (Budescu & Bar-Hillel, 1993; Roberts, 2006 among others). This scoring scheme is summarized by the following formula:

$$S = \begin{cases} 1 & \text{if the choice is right} \\ -\frac{1}{k-1} & \text{if the choice is wrong} \\ 0 & \text{if no option is chosen} \end{cases}$$

where k is the number of choices of an item.

A NM scoring scheme has to follow two rules (Roberts, 2006; Scharf & Baldwin, 2007)

- a pure random guess must have an expected score of 0,

- an informed guess (i.e. a guess after some choices have been correctly rejected) must have an expected score greater than 0 and less than the overall score of the item.

Betts et al. (2009) provide evidence that there is no statistically significant difference between the scores of students that knew that NM would be used in grading their tests (and thus they did not use random guessing) and those that were unaware that the scoring scheme would be NM (and thus they used random guessing). The conclusion is that NM reduces the effect of guessing on the score of a test. Furthermore, Burton (2001) states that NM should be considered as a disincentive for guessing and examinees should be informed that such a grading scheme will be used. On the other hand, Downing (2003) points out that NR methods give scores that are strongly correlated with those of the NR method and in general, the effect of guessing on test scores is overestimated.

Traub, Hambleton & Singh (1969) suggest that instead of NM of a wrong choice, the omission of the answer should be credited. Their scoring scheme is described by the following formula:

$$S = \begin{cases} 1 & \text{if the choice is right} \\ 0 & \text{if the choice is wrong} \\ \frac{1}{k} & \text{if no option is chosen} \end{cases}$$

The similarity of this method with NM is that the expected score of a pure random guess is the same with the score of an omission. On the other hand, it has increased validity and greater acceptance among examinees (Budescu & Bar-Hillel, 1993).

Correction for Guessing methods are designed to reduce the effect of random guessing in test scores. In what they fail to succeed is assessing partial knowledge. In NM every wrong answer is classified as a random guess and penalized, although a wrong answer could have a multitude of reasons and not all guesses are random (Akeroyd, 1982). The methods that follow are dealing with the problem of giving partial credit to partial knowledge. Detailed descriptions and categorization of such methods can be found at DeFinetti (1965) and Abu-Sayf (1979)

2.1.3. The Answer Until Correct method (AUC)

The Answer Until Correct (AUC) method (Gillman & Ferry, 1972) is a form of testing with MCQ in which the examinee is informed instantly about the correctness of his/her choice and, in case of an incorrect choice, is permitted to choose one of the remaining options. This procedure is repeated until the correct option is selected. The item score is computed from the number of attempts needed to choose the correct option. In a MCQ item with k alternatives, the scoring rule would be:

$$S(k,a) = k - a$$

where a is the number of attempts done to choose the correct option.

Due to the immediate response that provides to the examinee, this method is also referred to as Immediate Feedback Assessment Technique (Epstein et al., 2001; DiBattista, 2005)

Gillman & Ferry (1972) report that there was indication that AUC method can lead to increased internal reliability in comparison to the NR method. Epstein et al. (2001) point out that, although no significant difference in scores between the AUC and NR methods was observed, students tested with the AUC method tend to remember the correct options better than those examined with the NR scoring scheme. This result is an indication of the positive effect that immediate corrective feedback has on the learning process.

DiBattista (2005) summarizes the advantages of the AUC method:

- i.** promotes learning
- ii.** rewards partial knowledge with partial credit
- iii.** students have a strong preference compared to other methods of testing with MCQs

Implementations of the AUC method in paper utilize answer sheets that have the feedback of each option covered with a coating similar to that of scratch-off lottery tickets (DiBattista, 2005). Computer based implementation of the AUC method is obviously easier and more cost-effective.

2.1.4. Elimination Testing (ET)

Elimination Testing (ET) (Coombs, Millholand & Womer, 1956) is a method of testing using MCQs in which the student has to indicate the wrong options of an item than chose the correct one. For each correct elimination the

item score is credited with one point and in case of the elimination of the correct choice the penalty is $-(k - 1)$ points (k , the number of options of the item). For Multiple Choice Items with 4 options, the scoring formula is:

$$S = \begin{cases} 1 & \text{if one option is correctly eliminated} \\ 2 & \text{if two options are correctly eliminated} \\ 3 & \text{if three options are correctly eliminated} \\ -3 & \text{if the answer is incorrectly eliminated} \end{cases}$$

ET aims to reduce guessing and at the same time to give partial credit to partial knowledge. Criticism of the ET method centers upon the fact that by asking students to eliminate options promotes a negativistic approach to problem solving (Akeroyd, 1982).

Traub & Fischer (1977) report no statistically significant difference in score between NR and ET although they report a difference between Constructed Response Items and MCQs (both NR and ET). Bradbart et al. (2004) by comparing NR and ET, point out that the two methods produced similar reliability. Furthermore, using the ET method the partial knowledge had been assessed and guessing had been reduced. On the other hand, they point out the difficulty in constructing and scoring items with the ET method. Chang et al. (2007) by implementing ET on computers, deal with the difficulties of scoring and report results concerning partial knowledge and guessing similar to those in Bradbart et al. (2004).

Pollard (1989) suggests a combination of NR and ET in which the examinee has to either choose an option as correct or eliminate some options as incorrect. This method is presented with a scoring scheme that punishes random guesses with negative marking. Lau et al. (2011) present another combination of the NR and ET methods (NRET) implemented on computers.

2.1.5. Subset Selection Testing (SST)

Subset Selection Testing (SST) is a category of methods whose main characteristic is that the examinee is permitted to choose more than one options in a MCQ. The main scoring rule in an SST method is the following: the larger the number of options selected the lesser the score credited to the item.

An implementation of the SST method is suggested by Bush (2001) and referred to as “Liberal Tests”. According to that, in a Multiple Choice item with

k options the examinee is permitted to choose more than one option. If a is the number of choices and the answer is among them the item is credited with $\frac{k-a}{k-1}$ marks. If the answer of the item is not among the choices then the mark is $-\frac{a}{k-1}$. The following formula summarizes the scoring rule of “Liberal Tests”

$$S(k,a) = \begin{cases} \frac{k-a}{k-1} & \text{if the answer is among the choices} \\ -\frac{a}{k-1} & \text{if the answer is not among the choices} \end{cases}$$

where k is the number of options in a Multiple Choice Item and a is the number of options selected by the examinee.

Bush (2001) points out that “Liberal Tests” not only assess partial knowledge better than other methods of testing with MCQs but also give results of a higher resolution than conventional MC tests due to the greater number of alternative responses. In the same article, it is reported that the negative marking involved in the scoring made this kind of testing less attractive to the students. It is obvious that when $a=1$ (i.e. when the examinee chooses only one option) the scoring strategy is identical to NM. Frandsen και Schwartzbach (2006) suggest a variation of “Liberal Tests” with a more complex scoring rule.

Another method of the SST genre is “dual response” method (Akeroyd, 1982). In “dual response” the examinee is allowed to choose up to two options. In a MC item with 4 options the scoring scheme is:

$$S(k,a) = \begin{cases} \frac{1}{a} & \text{if the answer is among the choices} \\ 0 & \text{if the answer is not among the choices} \\ \frac{1}{4} & \text{if no option is chosen} \end{cases}$$

where a is the number of options selected ($a = 1$ or $a = 2$)

Akeroyd (1982) reports that “Dual response” gives more reliable scores since fluctuations due to random guessing are smoothed out and there is no

deflation of the scores like in Negative Marking. In addition, some information about random guesses and informed guesses is given to the examiners.

2.2. Students Assessment Preferences

Numerous studies in the past have investigated students' assessment preferences. The topic of discussion in many of them was students' preference between Multiple Choice Questions and Essay Type or Constructed Response Questions.

Zeidner (1990) reports that the majority of students believe that, in comparison to MCQ type exams, essay exams are more reflective of students' true achievements, are more suitable for assessing course objectives, and are generally fairer. In addition, students appear to be somewhat more anxious about MCQ than essay tests. However, students find MCQ type exams easier and less time consuming to prepare for, and feel less time-pressure when taking the exams in this format.

Birenbaum & Feldman (1998) report that students with good learning skills who have high confidence in their academic ability, tend to prefer the constructed-response type of assessment over the MCQ type. On the other hand, students with poor learning skills who tend to have low confidence in their academic ability prefer the MCQ over the constructed-response type of assessment. They also indicate that students with high test-anxiety tend to prefer the MCQ testing format.

Another aspect of the studies concerning students' assessment preferences is the investigation of the factors that affect them. A multitude of studies point out that, students' approaches to learning have a significant effect on assessment preference (Birenbaum, 1997; Furnham et al., 2008; Gijbels & Dochy, 2006). Furnham et al. (2008) report that students with a deep learning approach, favoured essay-type and oral exams as well as final dissertations. Gijbels & Dochy (2006) report that students with a deep approach to learning prefer assessment of high-order thinking tasks while those with a surface approach disfavor this type of assessment. Scouller (1998) and Struyven et al. (2005) claim that this relationship is two – way. Not only learning approaches can affect assessment preference but the type of test can make students adopt a learning approach.

Some reports on differences in assessment preference between the two genders can be found in the related literature (Ben-Chaim & Zoller, 1997; Birenbaum & Feldman, 1998). Birenbaum & Feldman (1998) point out that women prefer MCQ type testing over essay type questions while Ben-Chaim & Zoller (1997) suggest that women tend to prefer the more conventional types of examinations.

Chamorro-Premuzic et al. (2005) and Furnham et al. (2008) conclude that a significant factor that affects students' assessment preference is the personality of the student (as expressed by the Big Five categorization of personality traits). In both studies a negative relation between Openness and preference to MCQs is encountered (although only in the former this relation was statistically significant).

Not many studies have been encountered that deal with the question of students' attitude towards the alternative methods of MCQ type testing. DiBattista (2005) reports that the majority of students after been examined using the AUC method felt that it was fairer, easy to use and that the examination had a greater effect on learning than traditional methods.

3. Purpose of the Study

The purpose of the present study was to investigate students' attitudes towards the different methods of assessment that utilize Multiple Choice Questions. The main goal of this research was to provide evidence that would answer the following questions:

- what are the students' preferences and opinions regarding assessment using MCQs ?
- how do factors like gender, grade of education, school performance and course preference affect students' preference in a certain method ?
- what is the relation, if any, of student's preference in an assessment method and his/hers approach to learning ?

4. Method

4.1. Participants

178 students from the area of Thessaloniki took part in this study. There were 96 students of secondary education (between 15 and 18 years of age) and 82 university students. 100 of the participants were females and 78 males. Furthermore, from the secondary students' sample, data was collected concerning their preference in courses (mathematics, science and language) and their previous year final grades.

4.2. Questionnaires

4.2.1. Approaches to Learning

Secondary students' approaches to learning were measured by the Revised two-factor learning process questionnaire (R-LPQ-2F, Kember et al., 2004) while those of university students were measured by the Revised two-factor study process questionnaire (R-SPQ-2F, Biggs et al., 2001). Both of these questionnaires are used to classify students' approaches to learning in two major categories – surface and deep. The R-LPQ-2F consists of 22 questions and students' responses are recorded on a five-point Likert-type scale. It is designed to assess the learning approaches of students of secondary education. The R-SPQ-2F consists of 20 questions and students' responses are recorded on a five-point Likert-type scale. It is designed to assess the learning approaches of university students. Psychometric information regarding the reliability of the questionnaires is reported in Kember et al. (2004) and Biggs et al. (2001). Regarding our study, Cronbach's alphas are presented in Table 1. Such values are considered acceptable according to the creators of the questionnaires (Kember et al., 2004; Biggs et al., 2001) although they are just below the 0.70 cut-off point. Furthermore, in many studies, Cronbach's alphas between 0.65 and 0.70 are reported and considered acceptable.

Table 1: Cronbach's alpha of both scales of the R-LPQ-2F and R-SPQ-2F questionnaires

	Deep approach scale	Surface approach scale
R-LPQ-2F	0.694	0.663
R-SPQ-2F	0.665	0.698

4.2.2. Preferred Assessment Method

The participants were asked to state their opinion on six assessment and scoring methods, namely:

- Essay type questions.
- Multiple choice questions using the Number of Right scoring scheme (NR).
- Multiple choice questions using the Negative Marking method (NM) as described in the previous section.
- Multiple choice questions using the Answer Until Correct method (AUC).
- Multiple choice questions using the Elimination Testing (ET) method.
- Multiple choice questions using the Subset Selection Testing method (SST) as described in Bush (2001).

For this purpose, the participants were asked to respond to a 24-item questionnaire consisting of four items for each method. The four items intended to assess the perceived fairness, the perceived accuracy, the perceived student anxiety and the level of preference of each method. The four items were:

- *Indicate how fair you think is this method.*
- *Indicate how accurate you think is this method in measuring each person's level of knowledge.*
- *Indicate how stressful for the student you think this method is.*
- *Indicate the level of your personal preference in this method.*

A nine-point Likert-type scale was used to record the answer of each item.

4.3.Procedure

The students were given the questionnaires in class, under test conditions and given adequate time to think and respond. The assessment methods had been thoroughly explained to them in a previous meeting. In addition, a written explanation of the methods was given to them along with the questionnaires.

5. Results

5.1. Overall Results

Correlations between the four ratings (perceived fairness, perceived accuracy, perceived student anxiety and level of preference) have been computed for each method. Correlations between perceived fairness and level of preference were positive and significant at the 0.01 level: Essay Type ($r = 0.604$), NR ($r = 0.482$), NM ($r = 0.588$), AUC ($r = 0.449$), ET ($r = 0.384$) and SST ($r = 0.601$). Correlations between perceived accuracy and level of preference were also positive and significant at the 0.01 level though less strong in general: Essay Type ($r = 0.481$), NR ($r = 0.315$), NM ($r = 0.273$), AUC ($r = 0.336$), ET ($r = 0.475$) and SST ($r = 0.538$). Correlations between perceived student anxiety and level of preference were negative and not all of them significant: Essay Type ($r = -0.347$, *sig. at the 0.01 level*), NR ($r = -0.089$, *not significant*), NM ($r = -0.236$, *sig. at the 0.01 level*), AUC ($r = -0.055$, *not significant*), ET ($r = -0.153$, *sig. at the 0.05 level*) and SST ($r = -0.183$, *sig. at the 0.05 level*). Finally, correlations between perceived fairness and perceived accuracy were computed and found to be positive and significant at the 0.01 level: Essay Type ($r = 0.571$), NR ($r = 0.336$), NM ($r = 0.390$), AUC ($r = 0.451$), ET ($r = 0.474$) and SST ($r = 0.560$)

The mean rating scores in each item along with their standard deviations are presented in Table 2. The results clearly indicate a preference of the students towards MCQs graded with the Number of Right grading scheme. Their second preference is Essay Type questions while a strong disfavor for MCQs with Negative Marking is stated. The last three methods (Answer Until Correct, Elimination Testing and Subset Selection Testing) are preferred more than Negative Marking but less than Essay Type questions and Number of Right. Table 3 presents the pairwise comparison of the mean values of Level of Preference. The method used for the comparisons was the Repeated Measures ANOVA.

Table 2: Mean scores and SDs of preferred assessment methods

Method of Assessment		Level of Preference	Perceived Fairness	Perceived Accuracy	Perceived Student Anxiety
1. Essay Type	Mean	5.41	5.83	6.78	6.55
	SD	2.596	2.342	2.207	2.456
2. Number of Right	Mean	6.21	5.97	5.80	4.98
	SD	2.115	2.105	2.108	2.052
3. Negative Marking	Mean	3.20	3.50	4.47	6.73
	SD	2.195	2.391	2.410	2.095
4. Answer Until Correct	Mean	4.85	4.68	3.86	4.06
	SD	2.382	2.492	2.291	2.161
5. Elimination Testing	Mean	5.03	5.34	4.91	4.47
	SD	2.231	2.086	2.043	2.174
6. Subset Selection Testing	Mean	5.01	4.76	4.57	4.90
	SD	2.669	2.692	2.506	2.370

Regarding perceived fairness the results show the same ranking as in level of preference with a slight difference in the rankings of the AUC, ET and SST methods. As indicated in Table 2 the students seem to regard the ET method as fairer than AUC and SST methods.

Regarding perceived accuracy the results are quite different. The students seem to regard Essay Type questions as the most accurate method and the NR method as the second. The least accurate method in students' view is the AUC.

The most stressful method of assessment in student perception is Negative Marking, followed closely by Essay type questions. All other types of MCQ seem to be considered as less stressful. The smallest mean value is that of the AUC method.

Graph 1: Means of students attitudes towards the six methods of assessment

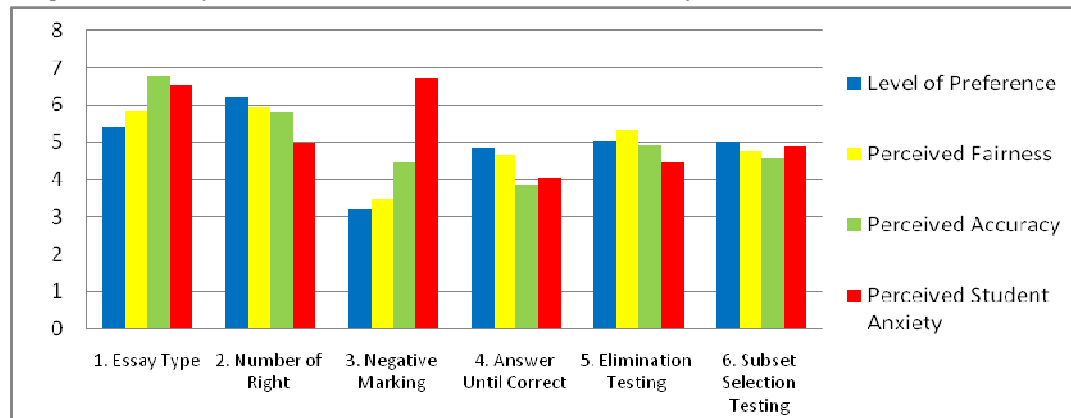


Table 3: Pairwise comparison of the mean values of Level of Preference using Repeated measures ANOVA

(I) Method of Assessment	(J) Method of Assessment	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	-,778*	,248	,030	-1,516	-,041
	3	2,188*	,264	,000	1,403	2,972
	4	,591	,300	,761	-,303	1,485
	5	,386	,293	1,000	-,485	1,257
	6	,398	,304	1,000	-,506	1,301
2	1	,778*	,248	,030	,041	1,516
	3	2,966*	,225	,000	2,295	3,637
	4	1,369*	,242	,000	,649	2,090
	5	1,165*	,243	,000	,441	1,889
	6	1,176*	,274	,000	,360	1,992
3	1	-2,188*	,264	,000	-2,972	-1,403
	2	-2,966*	,225	,000	-3,637	-2,295
	4	-1,597*	,223	,000	-2,260	-,933
	5	-1,801*	,223	,000	-2,464	-1,138
	6	-1,790*	,237	,000	-2,497	-1,083
4	1	-,591	,300	,761	-1,485	,303
	2	-1,369*	,242	,000	-2,090	-,649
	3	1,597*	,223	,000	,933	2,260
	5	-,205	,176	1,000	-,730	,320
	6	-,193	,238	1,000	-,901	,514
5	1	-,386	,293	1,000	-1,257	,485
	2	-1,165*	,243	,000	-1,889	-,441
	3	1,801*	,223	,000	1,138	2,464
	4	,205	,176	1,000	-,320	,730
	6	,011	,223	1,000	-,651	,674
6	1	-,398	,304	1,000	-1,301	,506
	2	-1,176*	,274	,000	-1,992	-,360
	3	1,790*	,237	,000	1,083	2,497
	4	,193	,238	1,000	-,514	,901
	5	-,011	,223	1,000	-,674	,651

Based on estimated marginal means

*. The mean difference is significant at the,05 level.

a. Adjustment for multiple comparisons: Bonferroni.

Table 4 presents the means of Level of Preference for each method after the sample has been categorized by gender. No significant difference between the means of the Male and Female categories is indicated in any of the methods, as the results of the independent samples T-tests showed.

Table 4: Means of Level of Preference categorized by gender

	Gender			
	Female		Male	
	Mean	Standard Deviation	Mean	Standard Deviation
Essay Type Level of Preference	5,29	2,552	5,56	2,661
NR Level of Preference	6,24	2,123	6,17	2,118
NM Level of Preference	3,14	2,211	3,28	2,186
AUC Level of Preference	4,87	2,355	4,82	2,432
ET Level of Preference	5,09	2,225	4,95	2,250
SST Level of Preference	4,94	2,585	5,09	2,788

Graph 2: Means of Level of Preference for the two genders

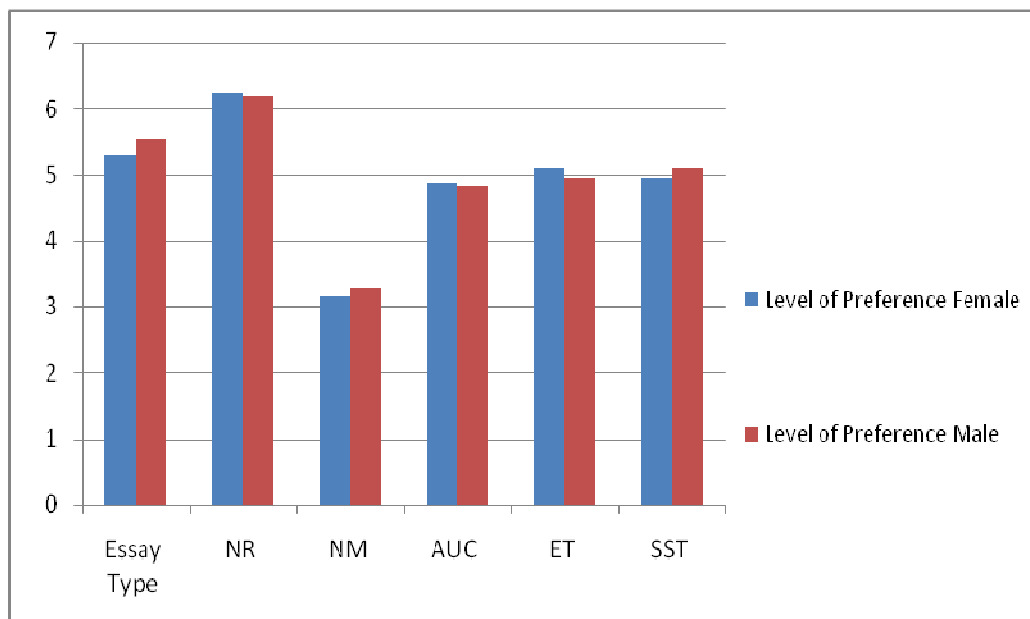


Table 5 presents the mean values of Level of Preference for each method after the sample has been categorized by grade. The difference between the means of the Secondary and University categories for the NR method is the only one significant at the 0.05 level according to the independent samples T-tests.

Table 5: Means of Level of Preference categorized by grade

	Grade			
	Secondary		University	
	Mean	Standard Deviation	Mean	Standard Deviation
Essay Type Level of Preference	5,22	2,653	5,63	2,526
NR Level of Preference *	5,86	2,019	6,62	2,165
NM Level of Preference	3,19	2,168	3,22	2,239
AUC Level of Preference	4,92	2,199	4,77	2,595
ET Level of Preference	5,14	2,208	4,90	2,264
SST Level of Preference	5,32	2,618	4,63	2,697

* the difference of mean values is significant at the 0.05 level

Graph 3: Means of Level of Preference of Secondary and University students

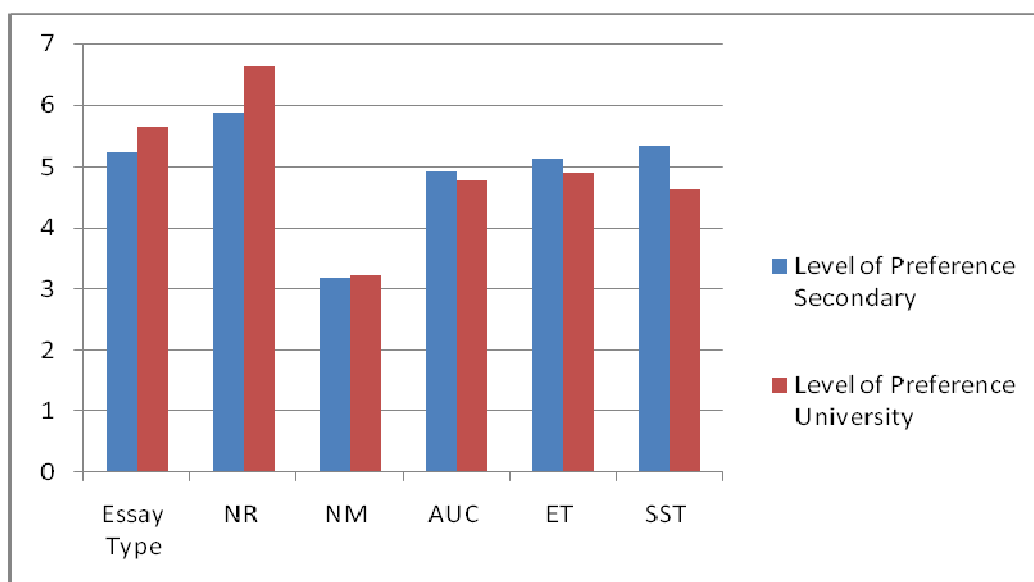


Table 6 presents the correlations between Level of Preference in the six methods. There are strong and significant negative correlations between Essay Type Level of Preference and the Level of Preference in AUC, ET and SST methods (correlations between Essay Type and AUC and Essay Type and ET are significant at the 0.01 level, while that between Essay Type and SST is significant at the 0.05 level). Correlations between Level of Preference in NM and AUC and Level of Preference in NM and SST are positive and significant (both at the 0.05 level). Correlations between Level of Preference in the AUC, ET and SST methods are strong, positive and significant (at the 0.01 level)

Table 6: Correlations between Level of Preference in different methods

	Essay Type Level of Preference	NR Level of Preference	NM Level of Preference	AUC Level of Preference	ET Level of Preference	SST Level of Preference
Essay Type Level of Preference	Pearson Correlation Sig. (2-tailed)	1				
NR Level of Preference	Pearson Correlation Sig. (2-tailed)	,036 ,631	1			
NM Level of Preference	Pearson Correlation Sig. (2-tailed)	-,067 ,373	,027 ,721	1		
AUC Level of Preference	Pearson Correlation Sig. (2-tailed)	-,310** ,000	-,025 ,740	,150* ,047	1	
ET Level of Preference	Pearson Correlation Sig. (2-tailed)	-,275** ,000	-,104 ,168	,093 ,215	,451** ,000	1
SST Level of Preference	Pearson Correlation Sig. (2-tailed)	-,158* ,036	-,143 ,057	,171* ,022	,200** ,008	,290** ,000

Prior to computing correlations between Level of Preference in each method and Approaches to Learning, the scale scores were converted in a percent scale in order for the scale scores of the two samples (university and secondary students) to be comparable.

Table 1Table 7 presents the correlations between Level of Preference in each method and Learning approach scale scores. Prior to computing correlations between Level of Preference in each method and Approaches to Learning, the scale scores were converted in a percent scale in order for the scale scores of the two samples (university and secondary students) to be comparable. There are positive and significant (at the 0.05 level) correlations between Level of Preference in Essay Type and Deep Approach scale score and Level of Preference in NR method and Deep Approach scale score. A negative and significant (at the 0.01) correlation between Level of Preference in NR method

and Surface Approach scale score was computed. Furthermore, a positive and significant (at the 0.05) correlation between Level of Preference in SST method and Surface Approach scale score was observed.

In addition, the difference between scale scores (Deep Approach scale score – Surface approach scale score) was computed. The results of correlations between Level of Preference in each method and Difference in Learning Approach scale scores, presented in Table 8, are similar to those of Table 7. Positive and significant (at the 0.05 and 0.01 level respectively) correlations were found between Level of Preference in Essay Type and Difference in Learning Approach scale scores and Level of Preference in NR method and Difference in Learning Approach scale scores.

Table 9 summarizes the means of Level of Preference in each method after the sample has been categorized by Learning Approach. If a students' Deep Approach scale score was higher than his/her Surface Approach scale score, the student is attributed a Deep Learning Approach (DA). Otherwise, the students' Learning Approach is considered to be a Surface Learning Approach (SA)

Table 7: Correlations between Level of Preference in each method and learning approach scale scores

		Essay Type Level of Preference	NR Level of Preference	NM Level of Preference	AUC Level of Preference	ET Level of Preference	SST Level of Preference
Deep Approach Scale Score (%)	Pearson Correlation	,185*	,190*	,026	-,033	,048	,122
	Sig. (2-tailed)	,013	,011	,730	,661	,524	,106
Surface Approach Scale Score (%)	Pearson Correlation	-,071	-,212**	-,074	,094	,004	,186*
	Sig. (2-tailed)	,343	,005	,327	,214	,961	,013

Table 8: Correlations between Level of Preference in each method and difference between learning approach scale scores

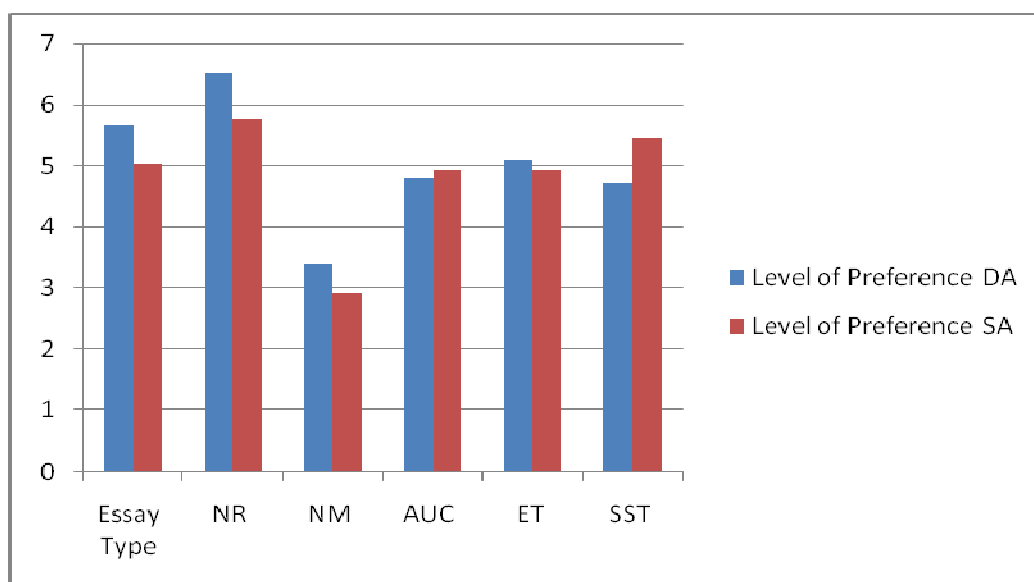
		Essay Type Level of Preference	NR Level of Preference	NM Level of Preference	AUC Level of Preference	ET Level of Preference	SST Level of Preference
Difference Between DA and SA scale scores	Pearson Correlation	,160*	,271**	,072	-,091	,025	-,073
	Sig. (2-tailed)	,032	,000	,343	,227	,745	,331

Table 9: Means of Level of Preference categorized by Learning Approach

	Learning Approach			
	DA		SA	
	Mean	Standard Deviation	Mean	Standard Deviation
Essay Type Level of Preference	5.65	2.657	5.04	2.475
NR Level of Preference *	6.51	1.992	5.76	2.226
NM Level of Preference	3.39	2.239	2.92	2.109
AUC Level of Preference	4.79	2.405	4.93	2.362
ET Level of Preference	5.09	2.276	4.93	2.173
SST Level of Preference	4.72	2.609	5.44	2.719

* the difference of mean values is significant at the 0.05 level

Graph 4: Means of Level of Preference of students with a Deep Learning Approach and Surface Learning Approach



In addition, factorial analysis on the six Levels of Preference was performed using the ANOVA method. The analysis indicated significant interactions between:

- the Learning Approach and Gender factors in Level of Preference in Essay Type (significant at the 0.01 level), AUC (significant at the 0.05 level) and SST (significant at the 0.05 level).

Table 10: ANOVA of Level of Preference in Essay Type Questions with Learning Approach and Gender factors

Tests of Between-Subjects Effects

Dependent Variable: Essay Type Level of Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	62,678 ^a	3	20,893	3,216	,024
Intercept	4727,720	1	4727,720	727,738	,000
LearnApp	23,334	1	23,334	3,592	,060
Gender	,024	1	,024	,004	,952
LearnApp * Gender	44,067	1	44,067	6,783	,010
Error	1130,384	174	6,496		
Total	6403,000	178			
Corrected Total	1193,062	177			

a. R Squared =,053 (Adjusted R Squared =,036)

Table 11: Mean values of Level of Preference in Essay Type Questions with the sample categorized by Learning Approach and Gender

Learning Approach * Gender

Dependent Variable: Essay Type Level of Preference

Learning Approach	Gender	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
DA	Female	5,172	,335	4,512	5,833
	Male	6,224	,364	5,506	6,943
SA	Female	5,452	,393	4,676	6,229
	Male	4,448	,473	3,514	5,382

Graph 5: Boxplots of Level of Preference in Essay Type Questions with the sample categorized by Learning Approach and Gender

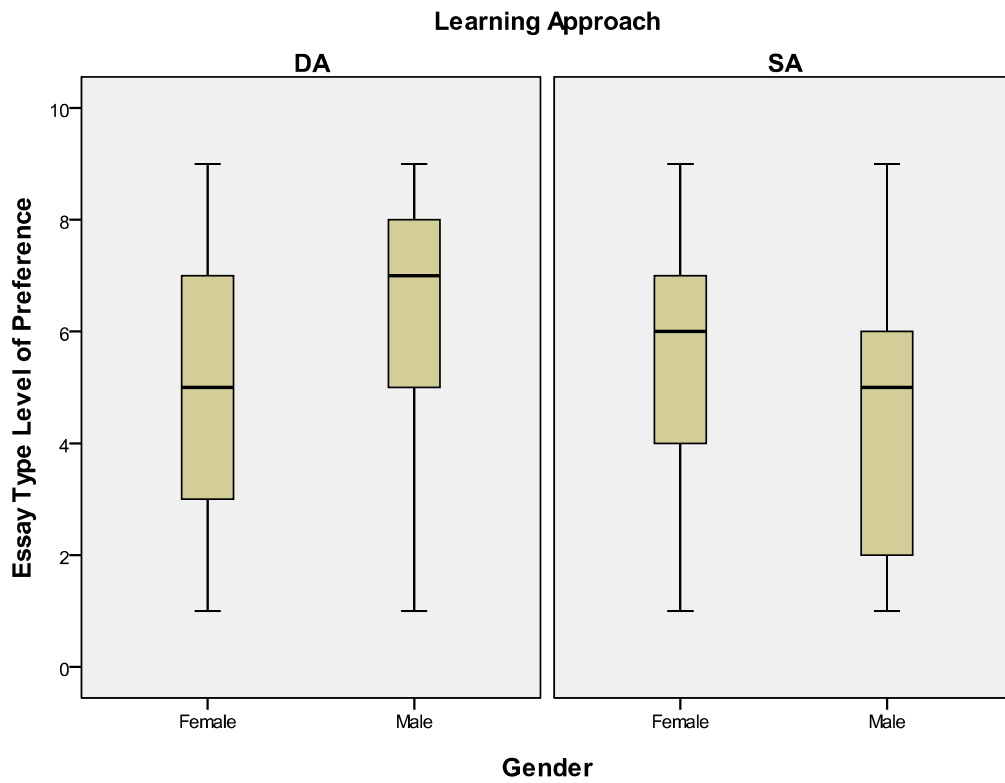


Table 12: ANOVA of Level of Preference in AUC with Learning Approach and Gender factors

Tests of Between-Subjects Effects

Dependent Variable:AUC Level of Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	24,739 ^a	3	8,246	1,465	,226
Intercept	3973,772	1	3973,772	705,711	,000
LearnApp	2,424	1	2,424	,430	,513
Gender	,589	1	,589	,105	,747
LearnApp * Gender	23,866	1	23,866	4,238	,041
Error	974,142	173	5,631		
Total	5158,000	177			
Corrected Total	998,881	176			

a. R Squared =,025 (Adjusted R Squared =,008)

Table 13: Mean values of Level of Preference in AUC with the sample categorized by Learning Approach and Gender

Learning Approach * Gender

Dependent Variable:AUC Level of Preference

Learning Approach	Gender	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
DA	Female	5,088	,314	4,467	5,708
	Male	4,449	,339	3,780	5,118
SA	Female	4,571	,366	3,849	5,294
	Male	5,448	,441	4,579	6,318

Graph 6: Boxplots of Level of Preference in AUC with the sample categorized by Learning Approach and Gender

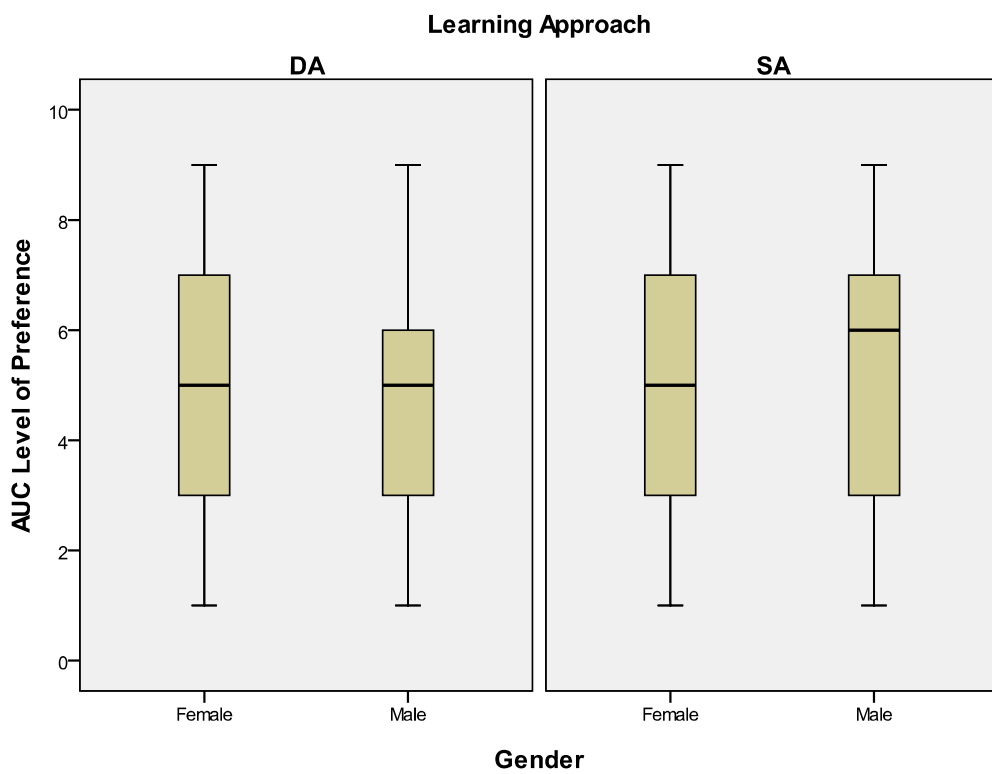


Table 14: ANOVA of Level of Preference in SST with Learning Approach and Gender factors

Tests of Between-Subjects Effects

Dependent Variable:SST Level of Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	58,686 ^a	3	19,562	2,831	,040
Intercept	4382,715	1	4382,715	634,273	,000
LearnApp	30,748	1	30,748	4,450	,036
Gender	6,103	1	6,103	,883	,349
LearnApp * Gender	35,253	1	35,253	5,102	,025
Error	1202,309	174	6,910		
Total	5721,000	178			
Corrected Total	1260,994	177			

a. R Squared =,047 (Adjusted R Squared =,030)

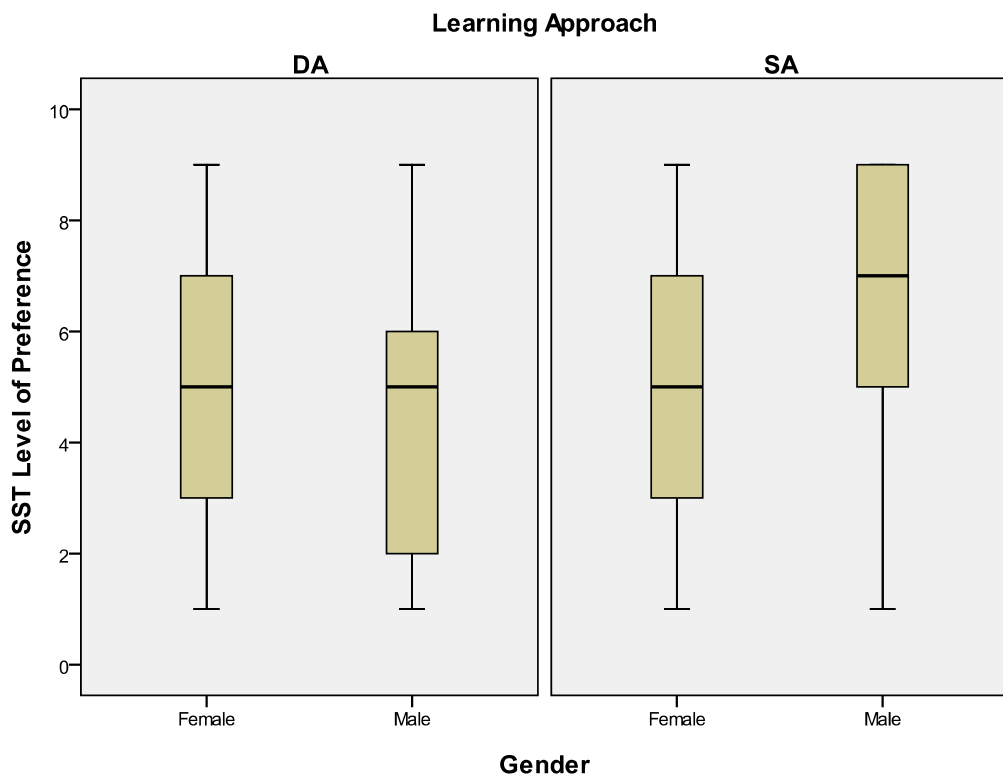
Table 15: Mean values of Level of Preference in SST with the sample categorized by Learning Approach and Gender

Learning Approach * Gender

Dependent Variable:SST Level of Preference

Learning Approach	Gender	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
DA	Female	4,966	,345	4,284	5,647
	Male	4,429	,376	3,687	5,170
SA	Female	4,905	,406	4,104	5,705
	Male	6,207	,488	5,243	7,170

Graph 7: Boxplots of Level of Preference in SST with the sample categorized by Learning Approach and Gender



- the Learning Approach and Grade factors in Level of Preference in SST (significant at the 0.05 level).

Table 16 : ANOVA of Level of Preference in SST with Learning Approach and Grade factors

Tests of Between-Subjects Effects

Dependent Variable: SST Level of Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	57,943 ^a	3	19,314	2,793	,042
Intercept	2987,353	1	2987,353	432,067	,000
LearnApp	1,693	1	1,693	,245	,621
Grade	21,237	1	21,237	3,072	,081
LearnApp * Grade	27,445	1	27,445	3,969	,048
Error	1203,052	174	6,914		
Total	5721,000	178			
Corrected Total	1260,994	177			

a. R Squared =,046 (Adjusted R Squared =,030)

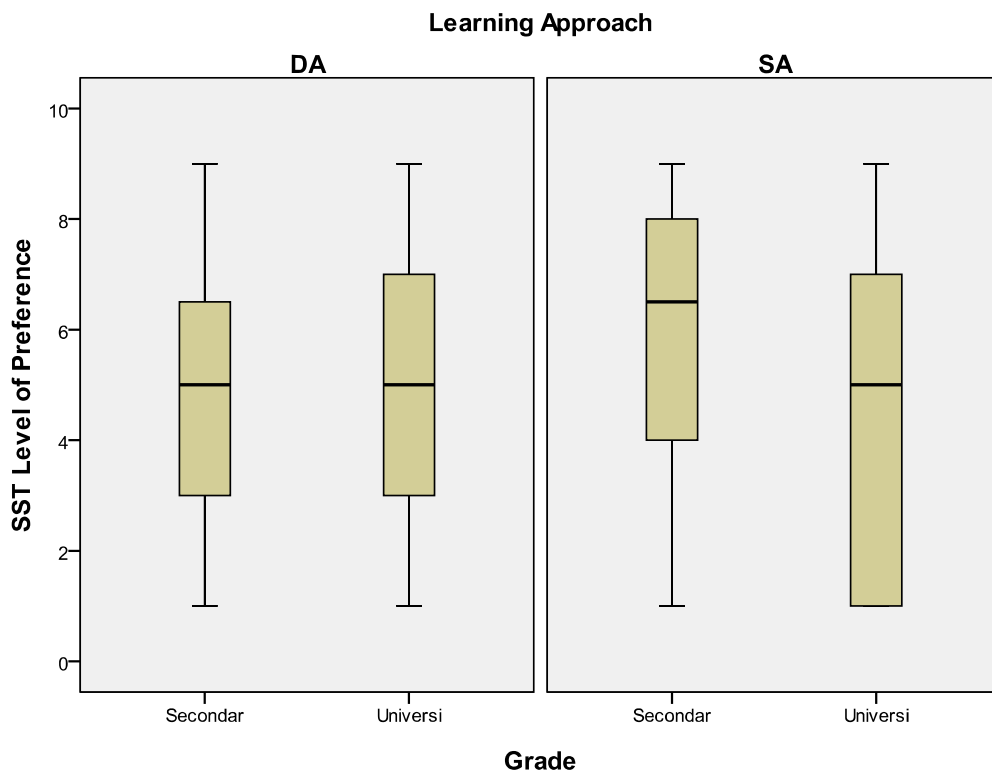
Table 17: Mean values of Level of Preference in SST with the sample categorized by Learning Approach and Grade

Learning Approach * Grade

Dependent Variable: SST Level of Preference

Learning Approach	Grade	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
DA	Secondary	4,650	,416	3,829	5,471
	University	4,761	,321	4,127	5,395
SA	Secondary	5,804	,351	5,110	6,497
	University	4,067	,679	2,727	5,407

Graph 8: Boxplots of Level of Preference in SST with the sample categorized by Learning Approach and Grade



- the Learning Approach, Gender and Grade factors in Level of Preference in NR and ET (both significant at the 0.05 level).

Table 18: ANOVA of Level of Preference in NR with Learning Approach, Gender and Grade factors

Tests of Between-Subjects Effects

Dependent Variable: NR Level of Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	71,435 ^a	7	10,205	2,409	,022
Intercept	3922,077	1	3922,077	925,960	,000
LearnApp	22,674	1	22,674	5,353	,022
Gender	4,693	1	4,693	1,108	,294
Grade	,146	1	,146	,035	,853
LearnApp * Gender	,005	1	,005	,001	,973
LearnApp * Grade	18,505	1	18,505	4,369	,038
Gender * Grade	8,114	1	8,114	1,916	,168
LearnApp * Gender * Grade	17,028	1	17,028	4,020	,047
Error	715,831	169	4,236		
Total	7611,000	177			
Corrected Total	787,266	176			

a. R Squared =,091 (Adjusted R Squared =,053)

Table 19: Mean values of Level of Preference in NR with the sample categorized by Learning Approach, Gender and Grade

Learning Approach * Gender * Grade

Dependent Variable:NR Level of Preference

Learning Approach	Gender	Grade	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DA	Female	Secondary	6,278	,485	5,320	7,235
		University	6,925	,325	6,283	7,567
	Male	Secondary	5,636	,439	4,770	6,503
		University	6,769	,404	5,972	7,566
SA	Female	Secondary	5,419	,370	4,690	6,149
		University	6,000	,621	4,775	7,225
	Male	Secondary	6,320	,412	5,507	7,133
		University	4,250	1,029	2,219	6,281

Table 20: ANOVA of Level of Preference in ET with Learning Approach, Gender and Grade factors

Tests of Between-Subjects Effects

Dependent Variable:ET Level of Preference

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	40,664 ^a	7	5,809	1,175	,319
Intercept	2855,724	1	2855,724	577,810	,000
LearnApp	,010	1	,010	,002	,965
Gender	2,051	1	2,051	,415	,520
Grade	,471	1	,471	,095	,758
LearnApp * Gender	2,946	1	2,946	,596	,441
LearnApp * Grade	3,830	1	3,830	,775	,380
Gender * Grade	,071	1	,071	,014	,905
LearnApp * Gender * Grade	27,121	1	27,121	5,487	,020
Error	840,195	170	4,942		
Total	5381,000	178			
Corrected Total	880,860	177			

a. R Squared =,046 (Adjusted R Squared =,007)

Table 21: Mean values of Level of Preference in ET with the sample categorized by Learning Approach, Gender and Grade

Learning Approach * Gender * Grade

Dependent Variable:ET Level of Preference

Learning Approach	Gender	Grade	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DA	Female	Secondary	4,889	,524	3,855	5,923
		University	5,325	,352	4,631	6,019
	Male	Secondary	5,773	,474	4,837	6,708
		University	4,333	,428	3,489	5,178
SA	Female	Secondary	5,161	,399	4,373	5,949
		University	4,364	,670	3,040	5,687
	Male	Secondary	4,720	,445	3,842	5,598
		University	6,000	1,112	3,806	8,194

Finally, six sets of regressions were performed. The depended variable was Level of Preference for each method and the independed variables were DA and SA scale scores, gender and grade. Not one of the models was significant at the 0.05 level. Stepwise regressions with depended variable the Level of Preference in each method did not produce any results that are worth to be presented.

5.2. Secondary Student Results

Table 22 presents the correlations between Level of Preference in each method and school performance. School performance is measured by the student's final grade of the previous year. There is a positive and significant (at the 0.01 level) correlation between Last year's final grade and Level of Preference in Essay Type questions, while a negative and significant (at the 0.05 level) correlation between Last Year's Final Grade and Level of Preference in the SST method was observed.

Table 22: Correlations between Level of Preference in each method and school performance

		Essay Type Level of Preference	NR Level of Preference	NM Level of Preference	AUC Level of Preference	ET Level of Preference	SST Level of Preference
Last Year's Final Grade	Pearson	,269**	,124	-,042	-,186	-,064	-,228*
	Correlation						
	Sig. (2-tailed)	,008	,227	,681	,069	,535	,025

The sample has been categorized in three ways. At Fist, by Preference in Mathematics (Means and SDs presented in Table 23), then by Preference in Science courses (Physics, Chemistry and Biology - Means and SDs presented in Table 24) and in the end, by Preference in Language Courses (Means and SDs presented in Table 25). For students who stated a preference in maths, the first choice of assessment method is MCQs with the NR grading scheme followed by Essay type questions. For the students who stated their disfavor in mathematics, the preferred assessment method is SST followed by NR and ET. The most preferred method among students who favor science courses is NR followed by SST, ET and AUC while among those who did not state a preference in science courses the choice is Essay type questions and NR. For students who stated their

liking for language courses the first choices were Essay Type Questions, ET and NR. On the other hand, among those who did not state a preference in language courses the most preferred method is NR, SST and AUC.

Table 23: Means of Level of Preference with the sample categorized by Preference in Mathematics

	Preference in Maths			
	No		Yes	
	Mean	Standard Deviation	Mean	Standard Deviation
Essay Type Level of Preference	4.79	2.710	5.56	2.582
NR Level of Preference	5.48	2.144	6.17	1.881
NM Level of Preference	3.02	2.247	3.31	2.118
AUC Level of Preference	4.57	2.120	5.19	2.241
ET Level of Preference	5.38	2.347	4.94	2.096
SST Level of Preference	5.76	2.555	4.98	2.639

Graph 9: Means of Level of Preference of secondary students categorized by preference in Mathematics

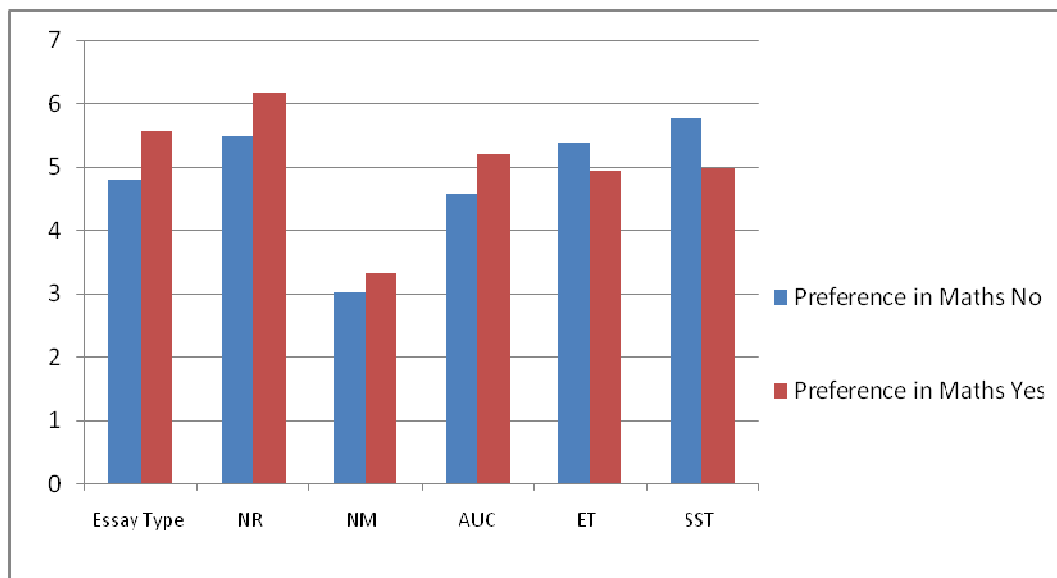


Table 24: Means of Level of Preference with the sample categorized by Preference in Science Courses

	Preference in Science			
	No		Yes	
	Mean	Standard Deviation	Mean	Standard Deviation
Essay Type Level of Preference	6.00	3.000	4.91	2.460
NR Level of Preference	5.59	1.986	5.97	2.036
NM Level of Preference	3.04	2.047	3.25	2.226
AUC Level of Preference	4.52	1.968	5.07	2.277
ET Level of Preference	5.07	2.433	5.16	2.133
SST Level of Preference	5.11	2.207	5.41	2.772

Graph 10: Means of Level of Preference of secondary students categorized by preference in Science Courses

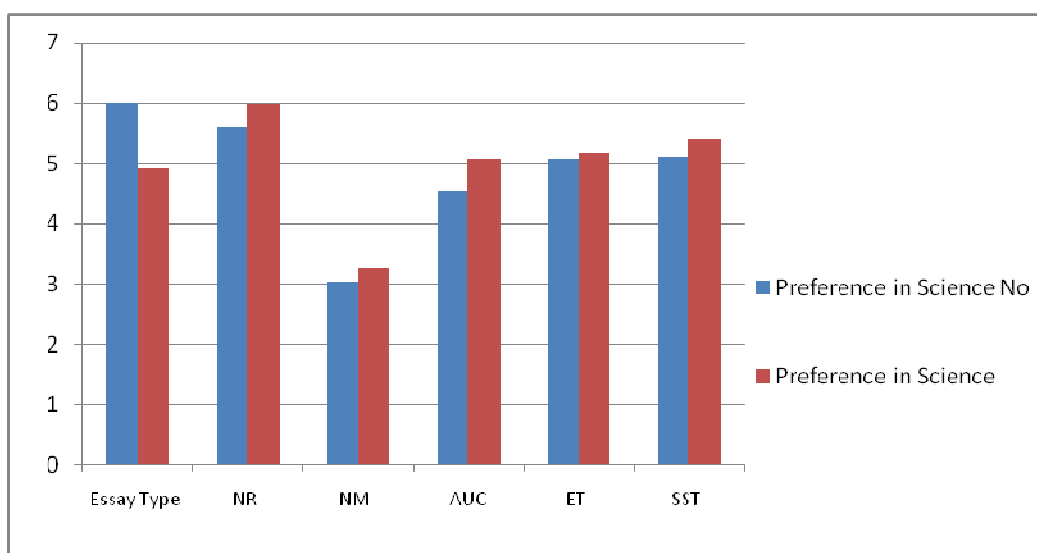
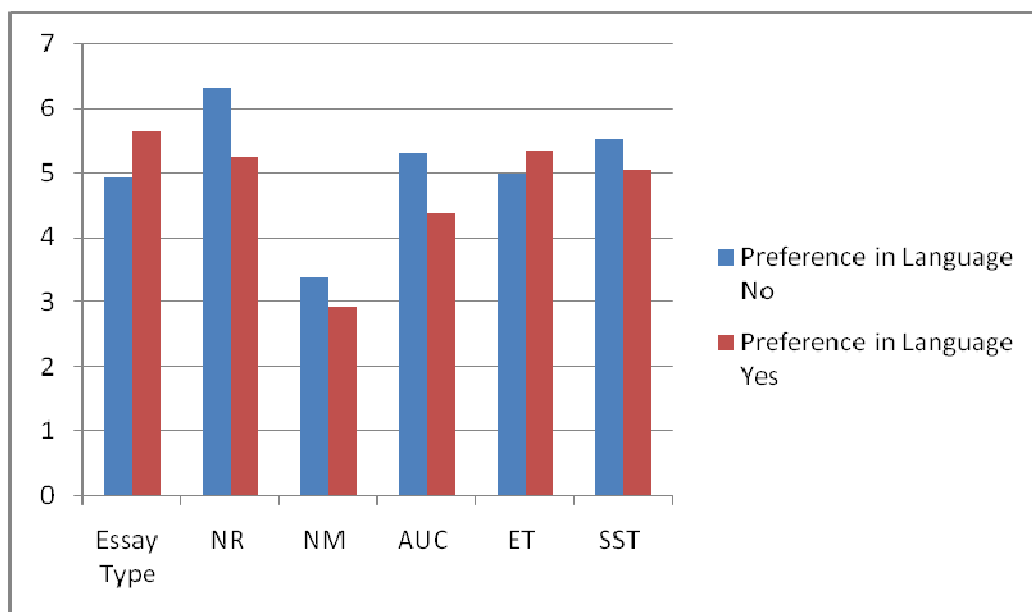


Table 25: Means of Level of Preference with the sample categorized by Preference in Language Courses

	Preference in Language			
	No		Yes	
	Mean	Standard Deviation	Mean	Standard Deviation
Essay Type Level of Preference	4.93	2.600	5.63	2.705
NR Level of Preference*	6.30	1.990	5.25	1.918
NM Level of Preference	3.38	2.245	2.93	2.055
AUC Level of Preference*	5.30	2.304	4.38	1.944
ET Level of Preference	4.98	2.153	5.35	2.293
SST Level of Preference	5.52	2.656	5.05	2.572

* the difference of mean values is significant at the 0.05 level

Graph 11: Means of Level of Preference of secondary students categorized by preference in Language Courses



In order to compare the mean values of Level of Preference between categories in course preference, independent samples T-tests were executed. The results showed significant differences only in some methods in the Preference in Language categorization of the sample (namely in the NR and AUC level of preference). However, some indication about the difference of the means in the other categorizations (Preference in Mathematics, Preference in Science) is observed though not supported by statistical evidence.

Finally, in six sets of regressions that were performed with the depended variable being Level of Preference for each method and the independed variables DA and SA scale scores, Preference in Maths, Preference in Science, Preference in Language and Gender, not one of the models had a significant fit. Stepwise regressions with depended variable the Level of Preference in each method did not produce any useful results also.

6. Conclusion

In students' perception of assessment, fairness and accuracy in measuring the knowledge of the examinee are closely related. Both of them seem to be a significant variable in shaping students' level of preference for a method of assessment. Less significant, yet existing, role seems to play the anxiety that a method brings to a student.

The students' most preferred method of the six proposed to them was NR, which is the conventional way of testing with MCQs. The least preferred method of assessment was, by far, MCQs with a negative marking scoring scheme. The second method in student preference was Essay Type questions but not significantly different than the remaining three (AUC, ET and SST). The mean values of each method in perceived fairness, perceived accuracy in measuring knowledge and perceived anxiety indicate that students preference in conventional MCQs is due to their perceived fairness and less anxiety. Although students regard Essay Type questions as more accurate in measuring knowledge, they seem to regard them as stressful. The sample's disfavor of negative marking is related to low mean values of perceived fairness and accuracy and a high mean value in perceived anxiety. Finally, worth mentioning are the relatively high mean values of perceived fairness of Elimination Testing although the level of preference in this method is not of the highest.

Indications of the relation between levels of preference in the six methods are presented. Students who have a higher level of preference in one of the alternative methods (AUC, ET and SST) tend to have high level of preference in the remaining two. Students that prefer Essay Type questions have a lower level of preference in alternative methods.

In this research, no evidence has been produced that gender by itself is a significant factor in forming the students' preference. Some implications that university level students have slightly different preferences than those of secondary education were presented, but the only one supported by statistical evidence is the difference in preference in conventional MCQs. (University students' level of preference in NR is significantly higher than that of secondary school students).

Another point of discussion in the present study is if and how the various approaches to learning that a student adopts influence his/her assessment preferences. The results indicate that the deeper the approach a student adopts the more he/she prefers conventional method of assessment as Essay Type and MCQs. Some significant correlations between Learning Approach scale scores and level of preference in each method support the former statement.

On the other hand, gender seems to be a significant factor in forming the level of preference when combined with other factors like Grade of Education or Learning Approach. The present study presents statistical evidence that:

- males with a deep approach to learning prefer Essay Type questions more than the rest of the students.
- males with a surface approach to learning prefer AUC and SST more than the rest of students.
- secondary school students with a surface approach to learning prefer SST more than the rest of the students.

In the part of the sample that consisted of secondary school students, correlational analysis presented indications that students with a high school performance tend to favor Essay Type Questions while those with a lower performance tend to favor the SST method. Implications that, students with a high performance have a lesser preference in the AUC and ET methods are not supported statistically. Moreover, the present study, presents hints that preference in certain courses (Mathematics, Science and Language) is related with a preference in some methods of assessment. However, the only statistically significant result is that students with a preference in language courses have a lower level of preference in conventional MCQs and AUC than the rest of the students.

The results of this study may be useful to educators that are interested in using MCQs in assessing student knowledge. Alternative MCQs address the problems of assessing partial knowledge and random guessing but student acceptance in a method of assessment is also important. Negative Marking seems to be rejected by the students as unfair and inconvenient method of assessment. The acceptance of the rest three methods (AUC, ET and SST) seems to depend on a multitude of factors like school performance, learning approach and course preference.

Further research is needed, in the direction of clarifying students' attitudes towards alternative MCQs. The lack of student experience with these methods might be the source of bias in their preferences. In order to explore the range, and finally, avoid that kind of bias, the study of students' attitudes before and after testing using each method seems to be a reasonable option.

7. References

- Abu-Sayf , F.K. (1979). Recent Developments in the Scoring of Multiple-Choice Items, *Educational Review*, 31:3, 269-279.
- Akeroyd, M. (1982). Progress in Multiple Choice Scoring Methods, 1977/81, *Journal of Further and Higher Education*, 6:3, 86-90.
- Ben-Chaim, D., & Zoller, U. (1997). Examination-type preferences of secondary school students and their teachers in the science disciplines, *Instructional Science* 25: 347–367, 1997.
- Betts, L. R., Elder, T. J., Hartley, J., & Trueman, M. (2009). Does correction for guessing reduce students' performance on multiple-choice examinations? Yes? No? Sometimes?, *Assessment & Evaluation in Higher Education*, 34:1, 1-15.
- Biggs, J., Kember, D., & Leung, D. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F, *British Journal of Educational Psychology*, 71, 133–149.
- Birenbaum, M., & Feldman, R.A. (1998). Relationships between learning patterns and attitudes towards two assessment formats, *Educational Research*, 40:1,90-98.
- Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations, *Higher Education* 33: 71–84.
- Bradbart, D., Parker D., & Stone, G. (2004). An Alternate Multiple Choice Scoring Procedure in a Macroeconomics Course, *Decision Sciences Journal of Innovative Education*, Volume 2, N.1.
- Budescu, D., & Bar-Hillel, M. (1993). To Guess or Not to Guess: A Decision-Theoretic View of Formula Scoring, *Journal of Educational Measurement*, Vol. 30, No. 4, pp. 277-291.
- Burton, R. F. (2005). Multiple-choice and true/false tests: myths and misapprehensions, *Assessment & Evaluation in Higher Education*, 30:1, 65-72.
- Burton, R.F. (2001). Quantifying the Effects of Chance in Multiple Choice and True/False Tests: Question selection and guessing of answers, *Assessment & Evaluation in Higher Education*, 26:1, 41-50.
- Bush, M. (2001). A Multiple Choice Test that Rewards Partial Knowledge, *Journal of Further and Higher Education*, 25:2, 157-163.
- Chamorro-Premuzic, T., Furnham, A., Dissou G., & Heaven, P. (2005). Personality and preference for academic assessment: A study with Australian University student, *Learning and Individual Differences* 15, 247–256.

- Chang, S.-H., Lin, P.-C., & Lin, Z. C. (2007). Measures of Partial Knowledge and Unexpected Responses in Multiple-Choice Tests. *Educational Technology & Society*, 10 (4), 95-109.
- Coombs, C.H., Millholand, J.E., & Womer, F.B. (1956). The assessment of partial knowledge, *Educational and Psychological Measurement*, 16, pp. 13-37.
- De Finetti, B. (1965). Methods for discriminating levels of partial knowledge. *British Journal of Mathematics and Statistical Psychology*, 18, 87-123.
- DiBattista, D. (2005). The Immediate Feedback Assessment Technique : A Learner-centered Multiple-choice Response form, *The Canadian Journal of Higher Education*, V.35, No. 4, p. 111 – 131.
- DiBattista, D., & Kurzawa, L. (2011). Examination of the Quality of Multiple-choice Items on Classroom Tests, *The Canadian Journal for the Scholarship of Teaching and Learning* Vol. 2: Iss. 2, Article 4.
- Downing, S.M. (2003). Guessing on selected-response examinations, *Medical Education* 37:670–671.
- Epstein, M., Epstein B., & Brosvic, G. (2001). Immediate Feedback During Academic Testing, *Psychological Reports* 88, p. 889 – 894.
- Frandsen, G., & Schwartzbach, M. (2006). A Singular Choice for Multiple Choice . *The SIGCSE Bulletin* Volume 38, Number 4.
- Furnham, A., Christopher, A., Garwood, J., & Martin, N. G. (2008). Ability, demography, learning style, and personality trait correlates of student preference for assessment method, *Educational Psychology: An International Journal of Experimental Educational Psychology*, 28:1, 15-27.
- Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference?, *Educational Studies*, 32:4, 399-409.
- Gilman, D.A., & Ferry, P. (1972). Increasing test reliability test through self-scoring procedures. *Journal of educational Measurement*, 9, 205-207.
- Haladyna, T.M., & Downing, S.M. (1989). A taxonomy of multiple-choice item writing rules. *Applied Measurement in Education*, 2(1), 37–50.
- Kember, D., Biggs J., & Leung, D. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire, *British Journal of Educational Psychology*, 74, 261–280.
- Lau, P.N.K., Lau, S.H., Hong, K. S., & Usop, H. (2011). Guessing, Partial Knowledge, and Misconceptions in Multiple-Choice Tests. *Educational Technology & Society*, 14 (4), 99–110.

- Pollard, G.H. (1989). Scoring to remove guessing in multiple choice examinations, *International Journal of Mathematical Education in Science and Technology*, 20:3, 429-432.
- Roberts, T.S. (2006). The use of multiple choice tests for formative and summative assessment. In *ACE 2006. Australian Computer Society*.
- Roediger, H.L., & Marsh, E.J. (2005). The positive and negative consequences of multiple choice testing. *Journal of Experimental Psychology*, 31: 1155_9.
- Scharf, E.M., & Baldwin, L.P. (2007). Assessing multiple choice question (MCQ) tests - a mathematical perspective, *Active Learning in Higher Education* , 8: 31.
- Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay, *Higher Education* 35: 453-472, 1998.
- Shaftel, J., & Shaftel, T.L. (2007). Educational assessment and the AACSB. *Issues in Accounting Education*, 22: 215-232.
- Siegfried, J.J., Saunders, P., Stinar, E., & Zhang, H. (1996). How is introductory economics taught in America? *Economic Inquiry*, 34, 182-192.
- Simkin, M.G., Kuechler, W.L. (2005). Multiple-Choice Tests and Student Understanding: What Is the Connection?, *Decision Sciences Journal of Innovative Education*, V.3 no 1.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review, *Assessment & Evaluation in Higher Education* Vol. 30, No. 4, pp. 331-347.
- Traub, R.E., Hambleton, R.K., & Singh B. (1969). Effects of Promised Reward and Threatened Penalty on Performance of a Multiple-Choice Vocabulary Test, *Educational and Psychological Measurement* 29: 847-861.
- Traub, R.E., & Fischer, C.W. (1977). On the Equivalence of Constructed-Response and Multiple-Choice Tests, *Applied Psychological Measurement*, Vol 1, No. 3, p 355 - 369.
- Trow, W.C. (1923). The psychology of confidence. *Archives of psychology*, 65, 1-47.
- Williams, J.B. (2006). Assertion-reason multiple-choice testing as a tool for deep learning: a qualitative analysis, *Assessment & Evaluation in Higher Education*, 31:3,287-301.
- Zeidner, M. (1990). College students' reactions towards key facets of classroom testing, *Assessment & Evaluation in Higher Education*, 15:2, 151-169.

APPENDIX

Questionnaire Used for Secondary Students

ΜΕΡΟΣ 1ο

Φύλλο:	A O	K O
--------	--------	--------

Τάξη:	A O	B O	Γ O
-------	--------	--------	--------

Κατεύθυνση:	Θετική O	Τεχνολογική O	Θεωρητική O
-------------	-------------	------------------	----------------

Μαθήματα που προτιμάτε	Μαθηματικά O	Φυσική O	Αρχαία Ελληνικά O	Νέα Ελληνικά O	Χημεία O	Βιολογία O	Ιστορία O	Οικονομία O	Άλλο O
------------------------	-----------------	-------------	----------------------	-------------------	-------------	---------------	--------------	----------------	-----------

Επίδοση (Τελική Επίδοση Προηγούμενης Χρονιάς):	<10 O	10-11 O	11-12 O	12-13 O	13-14 O	14-15 O	15-16 O	16-17 O	17-18 O	18-19 O	19-20 O
--	----------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------

ΜΕΡΟΣ 2ο

Παρακάτω υπάρχουν κάποιες ερωτήσεις για τις απόψεις σας για την εκπαίδευσή σας και για τον τρόπο που μελετάτε. Διαβάστε καθεμία ερώτηση και απαντήστε για το πόσο σας αντιπροσωπεύει

	δεν είναι ποτέ αλήθεια	είναι αλήθεια μόνο μερικές φορές	είναι αλήθεια περίπου τις μισές φορές	είναι αλήθεια αρκετές φορές	είναι αλήθεια σχεδόν πάντα
1. Πιστεύω ότι κάποιες φορές η μελέτη μου δίνει ικανοποίηση και χαρά	1 O	2 O	3 O	4 O	5 O
2. Προσπαθώ να συνδέω αυτά που έχω μάθει σε κάποιο μάθημα με αυτά που μαθαίνω σε άλλα μαθήματα	1 O	2 O	3 O	4 O	5 O
3. Απογοητεύομαι από έναν κακό βαθμό σε ένα διαγώνισμα και ανησυχώ για το πώς θα τα πάω στο επόμενο	1 O	2 O	3 O	4 O	5 O
4. Δεν βρίσκω κανένα σκοπό στο να μαθαίνω πράγματα τα οποία δεν είναι πιθανό να ερωτηθούν σε ένα διαγώνισμα	1 O	2 O	3 O	4 O	5 O
5. Αισθάνομαι ότι οποιοδήποτε θέμα μπορεί να είναι ενδιαφέρον αν αρχίσω να το καταλαβαίνω	1 O	2 O	3 O	4 O	5 O
6. Μου αρέσει να κατασκευάζω θεωρίες για να ταιριάξω πράγματα ξένα μεταξύ τους	1 O	2 O	3 O	4 O	5 O
7. Ακόμη και όταν έχω προετοιμαστεί πολύ καλά για ένα διαγώνισμα έχω ανησυχία για το ότι μπορεί να μην τα πάω καλά	1 O	2 O	3 O	4 O	5 O

8. Όταν αισθάνομαι ότι θα περάσω στην εξέταση, αφιερώνω όσο λιγότερο χρόνο γίνεται στη μελέτη του μαθήματος	1 0	2 0	3 0	4 0	5 0
9. Μελετώ πολύ γιατί βρίσκω το αντικείμενο ενδιαφέρον	1 0	2 0	3 0	4 0	5 0
10. Προσπαθώ να συνδέω την ύλη που μελετώ με αυτά που ήδη γνωρίζω για το αντικείμενο	1 0	2 0	3 0	4 0	5 0
11. Είτε μου αρέσει είτε όχι, βλέπω ότι η καλή σχολική επίδοση είναι σημαντικός παράγοντας επαγγελματικής επιτυχίας	1 0	2 0	3 0	4 0	5 0
12. Συνήθως περιορίζω τη μελέτη μου σε αυτά που έχουν οριστεί γιατί πιστεύω ότι δεν χρειάζεται να κάνω κάτι παραπάνω	1 0	2 0	3 0	4 0	5 0
13. Χρησιμοποιώ αρκετό από τον ελεύθερο χρόνο μου για να μάθω περισσότερα για ενδιαφέροντα θέματα που συζητήθηκαν σε κάποια μαθήματα	1 0	2 0	3 0	4 0	5 0
14. Όταν διαβάζω ένα σχολικό βιβλίο προσπαθώ να καταλάβω τι εννοεί ο συγγραφέας	1 0	2 0	3 0	4 0	5 0
15. Θέλω να αποφοιτήσω από το λύκειο γιατί πιστεύω ότι θα έχω καλύτερες προοπτικές επαγγελματικής αποκατάστασης	1 0	2 0	3 0	4 0	5 0
16. Πιστεύω ότι δεν χρειάζεται να μελετώ τα μαθήματα σε βάθος. Δεν είναι απαραίτητο να ξέρει κανείς πολλά για να τα καταφέρει στα περισσότερα μαθήματα	1 0	2 0	3 0	4 0	5 0
17. Στα περισσότερα μαθήματα έχω ερωτήσεις που θέλω να απαντηθούν	1 0	2 0	3 0	4 0	5 0
18. Μαθαίνω κάποια πράγματα «από έξω», επαναλαμβάνοντάς τα πολλές φορές μέχρι να τα μάθω, ακόμη και αν δεν τα καταλαβαίνω	1 0	2 0	3 0	4 0	5 0
19. Διαπιστώνω ότι πολλές φορές σκέφτομαι αυτά που μελέτησα σε διάφορες στιγμές όπως πχ όταν περπατάω, στο λεωφορείο, ξαπλωμένος κλπ	1 0	2 0	3 0	4 0	5 0
20. Νομίζω ότι ο καλύτερος τρόπος για να επιτύχω σε διαγώνισμα ή εξετάσεις είναι να θυμάμαι τις απαντήσεις από πολύ πιθανές ερωτήσεις	1 0	2 0	3 0	4 0	5 0
21. Μου αρέσει να κάνω αρκετή μελέτη για ένα θέμα έτσι ώστε να μπορώ να έχω τα δικά μου συμπεράσματα	1 0	2 0	3 0	4 0	5 0
22. Νομίζω ότι τα καταφέρνω καλύτερα σε διαγωνίσματα ή εξετάσεις όταν αποστηθίζω βασικά σημεία παρά όταν προσπαθώ να τα καταλάβω	1 0	2 0	3 0	4 0	5 0

ΜΕΡΟΣ 3ο

Παρακάτω περιγράφονται κάποιες μέθοδοι γραπτής αξιολόγησης. Διαβάστε προσεκτικά την περιγραφή κάθε μεθόδου και στη συνέχεια απαντήστε στις ερωτήσεις που ακολουθούν

(Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης). Ένα τεστ αποτελείται από 3 ερωτήσεις πλήρους ανάπτυξης στις οποίες ο εξεταζόμενος καλείται να απαντήσει αναπτύσσοντας το θέμα που ζητείται.

(Μέθοδος Καταμέτρησης Των Σωστών). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει μία επιλογή ή να μην επιλέξει καμία. Αν η επιλογή είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος ή δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος Καταμέτρησης Των Σωστών με Αρνητική Βαθμολογία). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει μία επιλογή ή να μην επιλέξει καμία. Αν η επιλογή είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος η ερώτηση βαθμολογείται με -1 μονάδες (δηλαδή αφαιρείται μία μονάδα από τη συνολική βαθμολογία). Αν δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος της Απάντησης μέχρι του Σωστού). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει μία επιλογή ή να μην επιλέξει καμία. Αν η επιλογή είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος τότε ο εξεταζόμενος καλείται να επιλέξει μία από 3 τις επιλογές που απομένουν. Αν η απάντηση είναι σωστή η ερώτηση βαθμολογείται με 2 μονάδες. Αν όχι, ο εξεταζόμενος καλείται να επιλέξει μία από τις δύο επιλογές που απομένουν. Αν η απάντηση είναι σωστή η ερώτηση βαθμολογείται με 1 μονάδα. Αν όχι η ερώτηση βαθμολογείται με 0 μονάδες. Αν δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος Διαγραφής των Λανθασμένων). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να **διαγράψει** από 1 έως 3 επιλογές τις οποίες θεωρεί λανθασμένες. Αν διαγράψει 3 λανθασμένες επιλογές η ερώτηση βαθμολογείται με 4 μονάδες. Αν διαγράψει 2 λανθασμένες επιλογές η ερώτηση βαθμολογείται με 3 μονάδες. Αν διαγράψει 1 λανθασμένη επιλογή η ερώτηση βαθμολογείται με 1 μονάδα. Αν σε οποιαδήποτε περίπτωση διαγραφεί η σωστή επιλογή, η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος Πολλών Επιλογών). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει από 1 έως 3 επιλογές ή να μην επιλέξει καμία. Αν έχει κάνει μία επιλογή και είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος τότε η ερώτηση βαθμολογείται με -1. Αν επιλέξει δύο απαντήσεις και ανάμεσα σε αυτές είναι η σωστή τότε η ερώτηση βαθμολογείται με 2 μονάδες. Αν ανάμεσα στις επιλογές δεν είναι η σωστή τότε η ερώτηση βαθμολογείται με -2. Αν επιλέξει τρεις επιλογές και ανάμεσα σε αυτές είναι η σωστή τότε η ερώτηση βαθμολογείται με 1 βαθμό. Αν ανάμεσα στις επιλογές δεν είναι η σωστή τότε η ερώτηση βαθμολογείται με -3. Αν δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

1. Πόσο δίκαιη πιστεύετε ότι είναι καθεμία από τις μεθόδους που περιγράφονται παραπάνω;

	Καθόλου Δίκαιη		Λίγο Δίκαιη		Μέτρια Δίκαιη		Αρκετά Δίκαιη		Πολύ Δίκαιη
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Των Σωστών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος της Απάντησης μέχρι του Σωστού	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Διαγραφής των Λανθασμένων	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Πολλών Επιλογών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Πόσο ακριβής στη μέτρηση της γνώσης πιστεύετε ότι είναι καθεμία από τις μεθόδους που περιγράφονται παραπάνω;

	Καθόλου Ακριβής		Λίγο Ακριβής		Μέτρια Ακριβής		Αρκετά Ακριβής		Πολύ Ακριβής
	○	○	○	○	○	○	○	○	○
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	○	○	○	○	○	○	○	○	○
Μέθοδος Καταμέτρησης Των Σωστών	○	○	○	○	○	○	○	○	○
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	○	○	○	○	○	○	○	○	○
Μέθοδος της Απάντησης μέχρι του Σωστού	○	○	○	○	○	○	○	○	○
Μέθοδος Διαγραφής των Λανθασμένων	○	○	○	○	○	○	○	○	○
Μέθοδος Πολλών Επιλογών	○	○	○	○	○	○	○	○	○

3. Πόσο άγχος πιστεύετε ότι σας προκαλεί την ώρα της εξέτασης καθεμία από τις μεθόδους που περιγράφονται παραπάνω;

	Καθόλου Άγχος		Λίγο Άγχος		Μέτριο Άγχος		Αρκετό Άγχος		Πολύ Άγχος
	○	○	○	○	○	○	○	○	○
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	○	○	○	○	○	○	○	○	○
Μέθοδος Καταμέτρησης Των Σωστών	○	○	○	○	○	○	○	○	○
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	○	○	○	○	○	○	○	○	○
Μέθοδος της Απάντησης μέχρι του Σωστού	○	○	○	○	○	○	○	○	○
Μέθοδος Διαγραφής των Λανθασμένων	○	○	○	○	○	○	○	○	○
Μέθοδος Πολλών Επιλογών	○	○	○	○	○	○	○	○	○

4. Να αναφέρετε το βαθμό προτίμησης σας σε καθεμία από τις μεθόδους

	Καθόλου		Λίγο		Μέτρια		Αρκετά		Πολύ
	○	○	○	○	○	○	○	○	○
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	○	○	○	○	○	○	○	○	○
Μέθοδος Καταμέτρησης Των Σωστών	○	○	○	○	○	○	○	○	○
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	○	○	○	○	○	○	○	○	○
Μέθοδος της Απάντησης μέχρι του Σωστού	○	○	○	○	○	○	○	○	○
Μέθοδος Διαγραφής των Λανθασμένων	○	○	○	○	○	○	○	○	○
Μέθοδος Πολλών Επιλογών	○	○	○	○	○	○	○	○	○

Questionnaire Used for University Students

ΜΕΡΟΣ 1ο

Φύλλο:	A O	Γ O
--------	--------	--------

ΜΕΡΟΣ 2ο

Παρακάτω υπάρχουν κάποιες ερωτήσεις για τις απόψεις σας για την εκπαίδευσή σας και για τον τρόπο που μελετάτε. Διαβάστε καθεμία ερώτηση και απαντήστε για το πόσο σας αντιπροσωπεύει

	δεν είναι ποτέ αλήθεια	είναι αλήθεια μόνο μερικές φορές	είναι αλήθεια περίπου τις μισές φορές	είναι αλήθεια αρκετές φορές	είναι αλήθεια σχεδόν πάντα
23. Πιστεύω ότι κάποιες φορές η μελέτη μου δίνει ένα αίσθημα βαθιάς προσωπικής ικανοποίησης και χαράς	1 0	2 0	3 0	4 0	5 0
24. Πιστεύω ότι, για να είμαι ικανοποιημένος, πρέπει να εργαστώ αρκετά σε κάποιο θέμα ώστε να μπορώ να διαμορφώνω τα δικά μου συμπεράσματα	1 0	2 0	3 0	4 0	5 0
25. Ο σκοπός μου είναι να περάσω το μάθημα κάνοντας όσο το δυνατόν λιγότερη δουλειά	1 0	2 0	3 0	4 0	5 0
26. Μελετώ σοβαρά μόνο το υλικό που δίνεται στην τάξη στα πλαίσια του μαθήματος	1 0	2 0	3 0	4 0	5 0
27. Αισθάνομαι ότι οποιοδήποτε θέμα μπορεί να είναι ενδιαφέρον αν αρχίσω να το καταλαβαίνω	1 0	2 0	3 0	4 0	5 0
28. Βρίσκω τα περισσότερα νέα θέματα ενδιαφέροντα και συχνά διαθέτω επιπλέον χρόνο για να πληροφορηθώ για αυτά	1 0	2 0	3 0	4 0	5 0
29. Δεν βρίσκω τις σπουδές μου πολύ ενδιαφέρουσες και επομένως κάνω την ελάχιστη δυνατή προσπάθεια	1 0	2 0	3 0	4 0	5 0
30. Μαθαίνω κάποια πράγματα «από έξω», επαναλαμβάνοντάς τα πολλές φορές μέχρι να τα μάθω, ακόμη και αν δεν τα καταλαβαίνω	1 0	2 0	3 0	4 0	5 0
31. Νομίζω ότι η μελέτη ακαδημαϊκών θεμάτων μπορεί να είναι τόσο ενδιαφέρουσα όσο ένα καλό βιβλίο ή μία καλή ταινία	1 0	2 0	3 0	4 0	5 0
32. Δοκιμάζω τον εαυτό μου για σημαντικά θέματα μέχρι να τα αντιλαμβάνομαι πλήρως	1 0	2 0	3 0	4 0	5 0
33. Νομίζω ότι τα καταφέρνω καλύτερα σε διαγωνίσματα ή εξετάσεις όταν αποστηθίζω βασικά σημεία παρά όταν προσπαθώ να τα καταλάβω	1 0	2 0	3 0	4 0	5 0

34. Συνήθως περιορίζω τη μελέτη μου σε αυτά που έχουν οριστεί γιατί πιστεύω ότι δεν χρειάζεται να κάνω κάτι περισσότερο	1 0	2 0	3 0	4 0	5 0
35. Μελετάω πολύ γιατί βρίσκω την ύλη ενδιαφέρουσα	1 0	2 0	3 0	4 0	5 0
36. Χρησιμοποιώ αρκετό από τον ελεύθερο χρόνο μου για να μάθω περισσότερα για ενδιαφέροντα θέματα που συζητήθηκαν σε κάποια μαθήματα	1 0	2 0	3 0	4 0	5 0
37. Δεν νομίζω ότι είναι χρήσιμο να μελετάω τα θέματα σε βάθος. Μου δημιουργεί σύγχυση και χάνω χρόνο ενώ αυτό που χρειάζεται είναι μία επιφανειακή επαφή.	1 0	2 0	3 0	4 0	5 0
38. Πιστεύω ότι οι καθηγητές δεν θα έπρεπε να περιμένουν από τους φοιτητές τους να διαθέτουν πολύ χρόνο για να μελετούν ύλη που είναι γνωστό ότι δεν θα εξεταστεί	1 0	2 0	3 0	4 0	5 0
39. Στα περισσότερα μαθήματα έρχομαι με ερωτήσεις στο μυαλό μου που θέλω να απαντηθούν	1 0	2 0	3 0	4 0	5 0
40. Πιστεύω ότι πρέπει κανείς να διαβάζει τα περισσότερα συγγράμματα που συνοδεύουν ένα μάθημα.	1 0	2 0	3 0	4 0	5 0
41. Δεν βλέπω το λόγο να μαθαίνω ύλη η οποία δεν είναι πιθανό να εξεταστεί	1 0	2 0	3 0	4 0	5 0
42. Νομίζω ότι ο καλύτερος τρόπος για να επιτύχω σε εξετάσεις είναι να θυμάμαι τις απαντήσεις από πιθανές ερωτήσεις	1 0	2 0	3 0	4 0	5 0

ΜΕΡΟΣ 3ο

Παρακάτω περιγράφονται κάποιες μέθοδοι γραπτής αξιολόγησης. Διαβάστε προσεκτικά την περιγραφή κάθε μεθόδου και στη συνέχεια απαντήστε στις ερωτήσεις που ακολουθούν

(Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης). Ένα τεστ αποτελείται από 3 ερωτήσεις πλήρους ανάπτυξης στις οποίες ο εξεταζόμενος καλείται να απαντήσει αναπτύσσοντας το θέμα που ζητείται.

(Μέθοδος Καταμέτρησης Των Σωστών). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει μία επιλογή ή να μην επιλέξει καμία. Αν η επιλογή είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος ή δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος Καταμέτρησης Των Σωστών με Αρνητική Βαθμολογία). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει μία επιλογή ή να μην επιλέξει καμία. Αν η επιλογή είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος η ερώτηση βαθμολογείται με -1 μονάδες (δηλαδή αφαιρείται μία μονάδα από τη συνολική βαθμολογία). Αν δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος της Απάντησης μέχρι του Σωστού). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει μία επιλογή ή να μην επιλέξει καμία. Αν η επιλογή είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος τότε ο εξεταζόμενος καλείται να επιλέξει μία από 3 τις επιλογές που απομένουν. Αν η απάντηση είναι σωστή η ερώτηση βαθμολογείται με 2 μονάδες. Αν όχι, ο εξεταζόμενος καλείται να επιλέξει μία από τις δύο επιλογές που απομένουν. Αν η απάντηση είναι σωστή η ερώτηση βαθμολογείται με 1 μονάδα. Αν όχι η ερώτηση βαθμολογείται με 0 μονάδες. Αν δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος Διαγραφής των Λανθασμένων). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να **διαγράψει** από 1 έως 3 επιλογές τις οποίες θεωρεί λανθασμένες. Αν διαγράψει 3 λανθασμένες επιλογές η ερώτηση βαθμολογείται με 4 μονάδες. Αν διαγράψει 2 λανθασμένες επιλογές η ερώτηση βαθμολογείται με 3 μονάδες. Αν διαγράψει 1 λανθασμένη επιλογή η ερώτηση βαθμολογείται με 1 μονάδα. Αν σε οποιαδήποτε περίπτωση διαγραφεί η σωστή επιλογή, η ερώτηση βαθμολογείται με 0 μονάδες.

(Μέθοδος Πολλών Επιλογών). Ένα τεστ αποτελείται από 20 ερωτήσεις πολλαπλής επιλογής με 4 επιλογές η καθεμία. Σε κάθε ερώτηση ο εξεταζόμενος καλείται να επιλέξει από 1 έως 3 επιλογές ή να μην επιλέξει καμία. Αν έχει κάνει μία επιλογή και είναι σωστή η ερώτηση βαθμολογείται με 3 μονάδες. Αν η επιλογή είναι λάθος τότε η ερώτηση βαθμολογείται με -1. Αν επιλέξει δύο απαντήσεις και ανάμεσα σε αυτές είναι η σωστή τότε η ερώτηση βαθμολογείται με 2 μονάδες. Αν ανάμεσα στις επιλογές δεν είναι η σωστή τότε η ερώτηση βαθμολογείται με -2. Αν επιλέξει τρεις επιλογές και ανάμεσα σε αυτές είναι η σωστή τότε η ερώτηση βαθμολογείται με 1 βαθμό. Αν ανάμεσα στις επιλογές δεν είναι η σωστή τότε η ερώτηση βαθμολογείται με -3. Αν δεν επιλεγεί καμία απάντηση η ερώτηση βαθμολογείται με 0 μονάδες.

5. Πόσο δίκαιη πιστεύετε ότι είναι καθεμία από τις μεθόδους που περιγράφονται παραπάνω;

	Καθόλου Δίκαιη		Λίγο Δίκαιη		Μέτρια Δίκαιη		Αρκετά Δίκαιη		Πολύ Δίκαιη
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Των Σωστών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος της Απάντησης μέχρι του Σωστού	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Διαγραφής των Λανθασμένων	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Πολλών Επιλογών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Πόσο ακριβής στη μέτρηση της γνώσης πιστεύετε ότι είναι καθεμία από τις μεθόδους που περιγράφονται παραπάνω;

	Καθόλου Ακριβής		Λίγο Ακριβής		Μέτρια Ακριβής		Αρκετά Ακριβής		Πολύ Ακριβής
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Των Σωστών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος της Απάντησης μέχρι του Σωστού	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Διαγραφής των Λανθασμένων	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Πολλών Επιλογών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Πόσο άγχος πιστεύετε ότι σας προκαλεί την ώρα της εξέτασης καθεμία από τις μεθόδους που περιγράφονται παραπάνω;

	Καθόλου Άγχος		Λίγο Άγχος		Μέτριο Άγχος		Αρκετό Άγχος		Πολύ Άγχος
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Των Σωστών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος της Απάντησης μέχρι του Σωστού	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Διαγραφής των Λανθασμένων	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Πολλών Επιλογών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Να αναφέρετε το βαθμό προτίμησης σας σε καθεμία από τις μεθόδους

	Καθόλου		Λίγο		Μέτρια		Αρκετά		Πολύ
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Ερωτήσεων Πλήρους Ανάπτυξης	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Των Σωστών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Καταμέτρησης Σωστών με Αρνητική Βαθμολογία	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος της Απάντησης μέχρι του Σωστού	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Διαγραφής των Λανθασμένων	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Μέθοδος Πολλών Επιλογών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>