

THE STUDENTS' PERCEPTIONS: ESSAY VERSUS MULTIPLE-CHOICE TYPE EXAMS

Dogan Tozoglu, Musavver D. Tozoglu

College of Education, Instructional Systems Program, Florida State University, USA

Ahmet Gurses, Cetin Dogar

Atatürk University, Department of Chemistry Education, Turkey

Abstract. *The purpose of this study was to 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 towards multiple choice test format compared to essay type formats in terms of the most critical dimensions assessed. These findings suggest that student perceptions should be taken into consideration while planning and constructing classroom testing. Suggestions for future research and implications for having students' perceptions are also discussed.*

Keywords: *test attitudes, essay exams, multiple choice exams, test format.*

Introduction

Assessment plays very significant role in education. It is an important aspect of teaching and learning, and if done correctly can provide important information to both teachers and students. Assessment is defined as the collection, synthesis, and interpretation of information to aid the teacher in decision-making (Airasian, 1997). It refers to a related series of measures used to determine a complex attribute of an individual or group of individuals (Oosterhof, 2001).

Assessment is necessary in order to determine whether students have met the learning objectives and have acquired the desired learning outcomes. Any assessment should measure the learning outcomes given for the course, the unit, or the day.

Instructors or teachers should always consider the degree to which students are achieving the objectives of the curriculum after any type of instruction. For this reason, classroom teachers should choose an assessment format that maintains fidelity to the instructional objectives and its feasibility in terms of costs, time and effort (Huff, 1998). They must reflect on the types of assessments they are utilizing and ask themselves how much and what type of assessing should be done.

In recent years, many forms of assessments have been used to measure knowledge and achievement. The most commonly used kinds of assessments are multiple choices, true-false, matching, short answer, and essay tests. They have either been used as is or used in combination with newer styles of assessment such as performance tests.

The process of constructing a reliable and valid test is a challenging task. Constructing a good test item is a deliberate process; it demands an understanding of the objectives and contents being assessed, how the thinking process of the learner is utilized, the reading and vocabulary level of the examinees, and knowledge of test-taking factors like response styles and test sophistication (Hopkins, 1998).

Depending on the purpose of the evaluation and the objectives to be measured, the form of the test may vary in length, difficulty, and format. The selections of item types or formats should be based on the types of outcomes you are trying to assess. Certain item types such as true/false,

supplied response, and matching, work well for assessing lower-order outcomes (i.e., knowledge or comprehension goals), while other item types such as essays, performance assessments, and some multiple choice questions, are better for assessing higher-order outcomes (i.e., analysis, synthesis, or evaluation goals).

The multiple-choice format and essay tests are the most popular item formats used in educational testing (Oosterhof, 2001). Many group-administered standardized tests consist entirely of multiple-choice items. The multiple-choice format is also used extensively in classroom tests, from elementary to graduate schools. The essay items represent a very flexible test format. It can potentially measure any skill that can be assessed with other formats of written tests. An essay item is uniquely able to assess a student's ability to communicate ideas in writing.

The multiple choice items traditionally consists of a stem that describes a problem and series of options, or alternatives, each representing possible answers to the stem. Normally, one option is correct, with the remaining alternatives referred to as distracters or foils. On the other hand, essay items require students to generate and construct an answer to an open-ended item, allow direct measurement of behaviors specified in a performance objective, and require students to provide a response rather than select one from alternatives provided.

In general, measurement literature describes and discusses considerations for constructing test items, advantages and disadvantages of various test formats, validity or reliability issues, etc. In addition, the literature also addressed other relevant issues such as factors for choosing one item format over another in planning a classroom test. Conversely, as test takers, the perspectives of the students, which is one of the major factor when planning a classroom test, have been generally disregarded (Zeidner, 1987).

Educators have been interested in student attitudes because of their possible impact on learning, the most important outcome of instruction. Student attitudes towards a situation can tell the teacher a great deal about the impact of that situation on the learning process. Attitudes are related to how people perceive the situations in which they find themselves (Smith, 1982, Dusic, 1998). Unfortunately, except for a few studies in high school settings (e.g. Zeidner, 1987), the area of student attitudes, dispositions, and preferences with respect to varying tests in general, or test formats in particular has not been researched recently at college or university settings. In most studies, only one or two dimensions such as test anxiety or general attitude toward a test format have been investigated. The researchers generally do not pay much attention to the examinees' perspective, one of the most potentially useful sources of information about the subjective qualities of a test or its components. The perceptions of students on particular test formats such as, perceived interestingness, usefulness, trickiness, fairness, etc., have not been adequately addressed in research (Zeidner, 1987).

Important, useful, and reliable data can be obtained through student perceptions. Students can produce a main source of information regarding the development and implementation of specific feedback systems designed to study examinees' reactions towards various facets of the classroom test. In addition, students can provide useful information on course elements such as textbooks, tests, and homework (Peterson, 2000). Students are good sources of information because they know their own situation well, uniquely know how students think and feel, experience the test first hand, and directly benefit from sufficient assessment. Therefore, examinees are one of the best sources of information about the subjective qualities of a test. Test constructors should take into consideration of the examinees' test attitudes and dispositions when deciding upon test construction and administration policy (Nevo, 1985; Zeidner, 1987).

Given the limited research on the area of assessing examinees' attitudes varying facets of classroom testing, the major purpose of the current study is to compare and contrast the preferences, attitudes, and perceptions of student examinees with respect to two of the most popular test formats

currently in use, essay versus multiple-choice test formats. This study also attempts to contribute and build upon the area of the literature, test preferences of students, which to date has not been extensively researched.

Methodology

Participants

The participants of this study were 50 freshman students at a small university situated in eastern part of the Turkey. 70 questionnaires were given to students, and 50 returned them completed, therefore, the return rate was 70%. All students were majoring in science education and were science teacher candidates. There were 15 female and 35 male students participated in the study. All of the students were white.

Instrument and Procedure

A test attitude inventory developed by Zeidner (1987) was used in order to gather data on students' perceptions and attitudes towards multiple choices and essay test formats. Students were informed about the difference between essay and multiple-choice test formats before responding to the instrument. Participants were told that teachers would have a chance to improve their classroom testing and interested in students' reactions towards various aspects of the most commonly used test formats. Student responded the instrument anonymously with no time limitation.

The instrument was constructed as a Likert-type rating scale and composed of 30 Likert type items, on a five-point range. Students were asked to rate each two type of classroom tests, essay versus multiple choice, separately along the following 10 different dimensions: their experience with the test formats (5 = very much, 1 = not at all), success expectancy (5 = very high, 1 = very low), reflection of student knowledge (5 = very high, 1 = very low), perceived facility (5 = very easy, 1 = very difficult), feeling at ease with format (5 = feeling very much at ease, 1 = feeling very ill at ease), perceived complexity (5 = not complex at all, 1 = very complex), perceived clarity (5 = very clear, 1 = very unclear), judged trickiness (5 = not tricky at all, 1 = very tricky), perceived fairness (5 = very fair, 1 = not at all fair), degree of anxiety evoked (5 = minimal degree of anxiety evoked, 1 = high degree of anxiety evoked).

The instrument consisted of several questions, two to four questions for each dimension, designed to assess each of the dimensions. There were a total of 30 questions. Sample questions from the instrument are as follows: In order to assess experience with the test formats, one question asks the students rate, "When you consider the exams you have taken, would you rate your experience with each exam format?" and to assess reflection of student knowledge one question asks students to rate "When you consider each exam type, which exam format do you think is more reflective indicators of the students' knowledge?" In order to assess trickiness one question asks students rate, "When you think about essay and multiple choice type exams, which exam type has more complicated or confusing questions?"

Higher scores on the items indicate more favorable dispositions toward the test format under consideration. In the analysis, individual scales were averaged and equal weights were used to have a composite attitude scale.

The alpha reliabilities of the instrument in each case were 0.83 and 0.86 for multiple choices and essay type of classroom test formats respectively. These alpha values can be considered reasonably satisfactory for group comparison purposes. These values were 0.85 in the previous study (Zeidner, 1987).

Results

This study provided students a chance to express their perceptions towards two types of assessment formats. The means and standard deviations for the individual and total scores are presented in Table 1. As indicated in Table 1, the total mean scores and standard deviations for the essay and multiple choice type exams were 3.06 (SD=1.31) and 3.47 (SD=1.06) respectively and difference was significant, $t(49)=5.04$, $p<0.001$.

Table 1. Ratings of student perceptions on essay versus multiple choice: means and standard deviations.

Scale	Essay		Multiple choice		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t values</i>
Experience	4.04	0.99	3.26	1.01	-4.30***
Success	2.94	1.09	3.90	0.95	4.13***
Knowledge	3.86	1.34	3.26	1.05	-2.45*
Facility	2.30	0.91	3.54	0.93	6.50***
At ease	2.78	1.31	3.76	0.94	4.02***
Complexity	2.68	1.27	3.48	1.25	2.93**
Clarity	3.04	1.10	3.84	0.80	3.52**
Trickiness	2.74	1.16	3.44	1.20	2.51*
Fairness	3.04	1.29	3.72	0.90	2.83**
Anxiety	2.44	1.34	3.26	1.07	3.21**
Total	3.06	1.31	3.47	1.06	5.04***

Note: All scales ranged from 1 to 5. Higher scores are indication of more favorable test attitudes than lower scores. For t test, degree of freedom was 49.

* $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Moreover, as shown in Table 1, while multiple choice type exams were rated significantly higher than essay type exams on 8 out of the 10 items, essay type exams were rated significantly higher in only 2 items appearing in the inventory. More specifically, in terms of being facility, students viewed multiple choice type format ($M=3.54$, $SD=.93$) being significantly easier than the essay type ($M=2.30$, $SD=0.91$), $t(49) = 6.50$, $p<0.001$. 52% of the students judged multiple-choice exams to be very easy or easy, contrary to only about 8% similarly perceived the essay exams. Furthermore, the multiple-choice exam was rated to be less complex ($3.48>2.68$) and clearer

(3.84>3.04) than essay type exam. While 58% of the students viewed multiple choice type exams not being complex at all and very little complex, the percentage was 21 for the same scales in essay type.

Interestingly, students considered the multiple-choice exam less tricky (3.44>2.74) and fairer (3.72>3.04) when compared with essay exam. Inspection of the frequencies of the scales showed that 52% of the sample perceived multiple choice items being not tricky or not at all tricky in comparison to only about 28% similarly perceiving the essay items. In addition, whereas 64% of the students judged the multiple choice type exams as being fair or very fair, only about 48% felt similarly about essay exams.

In terms of motivational variables being assessed, such as success expectancy, feeling at ease with format, and degree of anxiety evoked, the multiple choice exams were rated significantly higher than essay exams while students are taking the exam. In particular, students viewed multiple-choice exams as eliciting higher success expectancy than essay exams with the means of 3.90 and 2.94 respectively. They also thought that multiple choice exams (M=3.26) are being less anxiety evoking than essay exams (M=2.44). Lastly, they considered that multiple-choice exams (M=3.76) made respondents feel more at ease than essay exams (M=2.78) while taking exam. Frequency distributions showed that 68% of the sample expected to receive high or very high scores on multiple choice type exams, compared to only about 32% on essay exams. 54% of the students rated that essay exams were anxiety evoking or very anxiety evoking, whereas only about 28% felt the same way for multiple-choice exams. Likewise, 44% of students had a tendency of feeling ill at ease towards essay formats in comparison to only about 10% similarly perceiving multiple-choice formats.

On the other hand, student reported that they mostly came across essay type format (M=4.04) in their regular classes when compared to multiple-choice format (M=3.26), with the frequency of 70% and 40% respectively. In addition, students believed that essay exams (M=3.86) are more reflective indicators of the student knowledge compared to multiple-choice exams (M=3.26). Frequency distributions of the scale showed that 70% of the students viewed essay type exams being very high or high reflective indicators of knowledge, whereas 40% was similarly perceiving the multiple choice exams.

In sum, data from the study reflected a more positive attitudinal tendency of students towards multiple choice compared with essay type exams in terms of the majority of dimensions assessed.

Discussion

The purpose of this study was to obtain students' perceptions on two types of test formats. The most important findings from the present study were that students showed positive attitudes and more favorable perceptions towards multiple choice type exams than essay type exams along most dimensions assessed. Previous studies (Zeidner, 1987; Choi, 1998; Birenbaum & Feldman, 1998) support our findings and indicated that students have a positive preference towards multiple choice type exams.

Unlike the previous studies, the sample of this study were selected from different subject area and setting and investigated different dimensions related to students' perceptions. As mentioned before there are a few studies that have investigated this area or only one or two dimensions (e.g. Choi, 1998) related to students' perceptions of test formats have been researched. The participants of this study were science students at a small university. These current findings not only replicated and re-examined the previous work of Zeidner (1987), but also added to existing body of knowledge by including a different student group in terms of age, gender, cultural background and educational setting. The findings of this study were in line with the previous studies.

Multiple choice tests are the most common and perhaps the best tool for objective measurement of knowledge, ability or achievement because of its objectivity, simplicity, and automatic scoring, as well as the possibility of modifying a test based on empirical evidence (Ben-Simon, Budescu, & Nevo, 1997). Students viewed the multiple choice test formats as being less complex, clearer, fairer, and easier. This can be explained by the nature of the multiple choice test formats. According to Ebel and Frisbie (1991) the form of the multiple choice item, with the stem asking or implying a direct question, provides a realistic, naturally appropriate setting for testing. There tends to be less indirectness and artifice in multiple choices than in some other test formats. Students often find multiple-choice questions less ambiguous than other format items. In addition, less rote memorization of factual material and somewhat less time and effort for adequate preparation are required while preparing multiple-choice exams normally than essay exams. There is also no need to express answers in written form.

One of the most interesting findings was that although students believed that multiple choice formats were fairer, they perceived that essay formats reflected students' knowledge better. A possible explanation for multiple-choice exams being perceived to be fairer may be due to the fact that probability of incorrectly scoring a multiple-choice item is less. That means responses to multiple-choice items are objectively scored. It is easier to justify the score to the student in multiple-choice exams. Furthermore, decisions concerning passing or failing are better founded (Kamps & Lint, 1975).

On the other hand, the reliability of essay tests usually suffers from inconsistencies in scoring students' responses (Oosterhof, 2001). It has been found that it is possible to see that different scorers often give different grades to the same essay response; the same scorer often assigns different grades to the same essay response on different occasions (Tuckman, 1993). In order to provide support to multiple-choice exams as an alternative to the traditional or essay exams, Holley & Jenkins (1993) mentioned that multiple choice testing is the best available testing methodology to achieve the desired qualities of reliability and validity in the assignment of grades based on tests to measure student learning.

Students believed that essay exams reflect student knowledge much better than multiple choices. One plausible explanation is that essay items allow students construct their own responses, require students to select and organize the content for discussion, develop a rationale to support their thinking and position taken, consider multiple perspectives and present their ideas logically (Oermann, 1999). Essay items also provide students the opportunity of accurately and optimally expressing their knowledge and ideas in writing. On the other hand, even though the probability of guessing the correct answer is less in multiple-choice exams, there is still a possibility for guessing. Partial information provided students by the availability of options could also provide a relevant clue to the correct answer.

Another explanation for this result would be that the participants in this study were more familiar with essay type exams than multiple choices. The reason for this could be teachers and instructors mostly prefer essay tests because it is easier to construct an essay test. Constructing an essay test may be considerably less labor intensive than constructing a multiple-choice test (Tuckman, 1993). Constructing multiple-choice tests and having a question item pool take much more time. Since students mentioned they came across essay type exams much more or had more experience, that could lead them to think essay type exams reflects students' knowledge better, although multiple choice exams can provide a more adequate sampling of the content, involve more items, and require less time to record responses. In the previous study conducted by Zeidner (1987), students (70% of his sample) believed that essay formats reflected students' knowledge better, which supported our findings. However, only 51% of his sample perceived multiple-choice formats fairer, which partially supported our findings (64% in our sample).

Students also believed that they have a better chance of succeeding and have less anxiety on multiple choices relative to the essay type exams. A possible explanation for higher expectancy is that multiple-choice exams are relatively easier than essay exams. Since options on multiple choice tests are made available to students and they can have an option to guess, they feel sense of security and more confidence while taking the test. Students also do not worry about the possibility of being unfairly scored, making spelling mistakes, having poor writing ability, and preparing and making too much effort to succeed. Essay exams, however, requires students having additional effort and emotional energy in order to select, organize, and express their ideas (Zeidner, 1987). In addition, there is no availability of information or clues leading the correct answer. Studies also support these findings and showed that multiple-choice tests had no apparent effect on test anxiety. The results of the study conducted by Choi (1998) clearly indicate that the essay test format is related to higher levels of test anxiety in college students than is multiple-choice test formats.

One of the major limitations of this study is that the sample was a relatively small, non-random sample, which was unbalanced in terms of gender, and the research conducted among science education students only. It is possible that different results could be obtained from different age groups, different educational and cultural settings. Further research needed to strength the reliability and validity of the findings.

Conclusion

Students are the most affected by the educational testing. Therefore, teachers and measurement specialists should take into consideration of students' attitudes and perceptions regarding test formats, because they are good sources of information about a test's face validity besides its content, construct, and predictive validity. It is important to have evidence for face validity of the tests developed by the teachers from the students' critical perspectives.

Feedback from the students on various components or facets of classroom tests is valuable source of information, because their perspectives affect test preparation behavior, student cooperation and test motivation during the exam, and influence the level of test performance and attainment on the exam (Zeidner, 1987). If students have a positive tendency toward a particular test format, the possibility of student cooperation, teacher-student rapport, and test motivation would be enhanced, while aversive emotional reactions and harmful motivational tempers would be lessened.

Since students had strong preference for multiple choice over essay type formats in this study, teachers should pay attention and give careful consideration and weight to the multiple choice format, when they are initially planning a classroom test and deciding appropriated item format in a previously planned test.

References

- Airasian, P.W. (1997). *Classroom Assesment* (3rd ed.). NewYork: McGraw-Hill.
- Ben-Simon, A., Budesco, D.V., 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 towards 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, 616-20.
- 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.

Ebel, R., Frisbie, D. A. (1991). *Essentials of Educational Measurement*. Englewood Cliffs, New Jersey: Prentice Hall.

Holley, J.H., Jenkins, E.K. (1993). The relationship between student learning style and performance on various test questions formats. *Journal of Education for Businesses*, 68, 5, 301-308.

Hopkins, K. (1998). *Educational and Psychological Measurement and Evaluation*. Needham Heights, MA: Allyn & Bacon.

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).

Kamps, H. J. L., Van Lint, J. H. (1975). A comparison of a classical calculus test with a similar multiple choice test. *Educational Studies in Mathematics*, 6, 3, 259-271.

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

Oermann, M. (1999). "Developing and scoring essay tests". *Nurse Educator*, 24, 2, 29-32.

Oosterhof, A. (2001). *Classroom Applications of Educational Measurement*. Upper Saddle River, New Jersey, Columbus, Ohio: Merrill Prentice Hall, Inc.

Peterson, K.D. (2000). *Teacher Evaluation: A comprehensive guide to new directions and practices* (2nd ed.). Thousand Oaks: Corwin Press, Inc.

Smith, M.J. (1982). *Persuasion and Human Action*. Wadsworth, CA: Belmont.

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 student's perspective". *Journal of Educational Research*, 80, 6, 352-58.

Резюме

ПЕРСПЕКТИВЫ СТУДЕНТОВ ЕСТЕСТВЕННИКОВ В ТУРЦИИ К ДВУМ ФОРМАМ ОЦЕНКИ: ЭССЕ ПО СРАВНЕНИЮ С АЛЬТЕРНАТИВОЙ

Доган Тозоглу, Мусаввер Д. Тозоглу, Ахмет Гурсес, Цетин Догар

Цель этого исследования состояла в том, чтобы исследовать восприятие студентов к эссе по сравнению с альтернативными формами экзаменов. Пятьдесят студентов естественников участвовали в этом исследовании.

Установлено, что студенты показали значительно ($p < 0.001$) более благоприятные отношения к альтернативной форме оценки, например тесту, по сравнению с формой оценки типа эссе. Эти полученные данные предлагают, чтобы студенческое восприятие было учтено, планируя и конструируя тесты и проводя само тестирование. В статье обсуждены предложения для будущего исследования. Одни из самых интересных полученных данных были то, что, студенты полагали, что альтернативные формы оценки были более справедливы. Возможное объяснение альтернативных экзаменов - вероятность неправильного выигрыша альтернативного пункта меньше. Одно из главных ограничений этого исследования то, что выборка была относительно мала, неслучайная, которая была неуравновешенна в отношении пола, и исследования, проводимого только среди студентов естественников. Возможно, что различные результаты могли быть получены от различных возрастных групп.

Ключевые слова: тестирование студентов, эссе, альтернативный тест, отношения студентов.

Dogan Tozoglu

College of Education, Instructional Systems Program, Florida State University,
Tallahassee, FL, 32306 USA
E-mail: dogantozoglu@yahoo.com

Musavver D. Tozoglu

College of Education, Instructional Systems Program, Florida State University,
Tallahassee, FL, 32306 USA

Ahmet Gurses

Department of Chemistry, K.K.Egitim Faculty, Ataturk University
25240 Erzurum, Turkey

Cetin Dogar

Department of Science Education, Erzincan Egitim Faculty, Ataturk University
24030 Erzurum, Turkey
E-mail: chemistcetin@yahoo.co.uk

Copyright of Journal of Baltic Science Education is the property of Scientific Methodical Center and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.