AN ALTERNATIVE WAY OF TESTING AND SCORING: MULTIPLE-ANSWER MULTIPLE-CHOICE QUESTIONS FOR CLASSROOM USE

Kürşat CESUR

ABSTRACT

Single-Answer Multiple-choice (SAMC) test technique is one of the most commonly used objective test techniques. Most of the teachers and testing organizations prefer SAMC questions as these tests provide "high score reliability, ease of administration and scoring, usefulness in testing varied content, and objective scoring" (Kurz, 1999:3). However, such questions have some disadvantages such as "decreased validity due to guessing and failure to credit partial knowledge" (Kurz, 1999:2). Many scholars provided suggestions to overcome such disadvantages. Different scoring methods for SAMC questions have been created so far. In this paper, it is claimed that using Multiple-answer Multiple-choice (MAMC) questions will also eliminate some of these disadvantages of SAMC questions. In addition, it is aimed to suggest a practical way of scoring MAMC questions. This study is significant as there is a very limited literature on the use of MAMC questions for classroom use. The practical way of scoring MAMC questions for classroom use will be an initial step and it will pave the way for the new developments in the use of MAMC questions in the language tests. This alternative way of testing and scoring MAMC questions is open to discussion. There is nothing best, but something better which will help the further developments in testing the examinees' success.

Key Words: *Testing, Objective Tests, Multiple-answer Multiple-choice Questions*

1. INTRODUCTION

Teaching and learning is a broad process. This process does not finish when the students have learned the subject. The students' success should also be evaluated and measured at the end of this process. In this respect, testing becomes indispensable part of the teaching and learning process as it helps teaching. For many years,

students' success has been evaluated using various test techniques. Of all these techniques, multiple-choice questions have been one of the most commonly used one by the institutions, educators, test designers, teachers, so and so forth. However, when a multiple-choice question is to be asked, it is usually supposed to have one correct answer. This is because single-answer multiple-choice (SAMC) questions are used much more frequently than the multiple-answer multiple-choice (MAMC) questions.

In this paper; after a brief description of objective and subjective scoring, the disadvantages of SAMC questions are discussed. Then, a practical way of using MAMC questions and scoring these questions are suggested by the author. Finally, a conclusion and implications for ELT are drawn.

1.1 Statement of the Problem

SAMC test technique is one of the most commonly used objective test techniques. Most of the teachers and testing organizations prefer SAMC questions as these tests provide "high score reliability, ease of administration and scoring, usefulness in testing varied content, and objective scoring" (Kurz, 1999:3). However, such questions have some disadvantages such as "decreased validity due to guessing and failure to credit partial knowledge" (Kurz, 1999:2). Many scholars provided suggestions to overcome such disadvantages. Different scoring methods for SAMC questions have been created so far. The author of this paper claims that using MAMC questions will also eliminate some of these disadvantages of SAMC questions.

1.2 Purpose and Significance of the Study

This paper aims to suggest a practical way of scoring MAMC questions. Many studies regarding the SAMC questions have been carried out so far. Nevertheless, this study is significant as there is a very limited literature on the use of MAMC questions for classroom use. The practical way of scoring MAMC questions for classroom use is an initial step and it will pave the way for the new developments in the use of MAMC questions in the language tests.

2. OBJECTIVE TEST TECHNIQUES

The terms *objective* and *subjective* are used for the scoring of tests. A test can be called as objective or subjective by determining the way how the teacher scores the students' performances. Subjectiveness can be used for everything. Not only the scoring but also the construction of tests can be subjective as the testers ask whatever they want to ask. However, "only the *scoring* of a test can be described as objective" (Heaton, 1988:25). To be called as an objective test, a test must "have only one correct answer (or, at least, a limited number of correct answers)" (Heaton, 1988:25).

Madsen (1983) gives many examples of the objective test techniques used in language testing. Like Madsen, many authors (Heaton, 1988; Genesee & Upshur, 1996; Hughes, 1989; Alderson, 2000; Weir, 1990) provide examples of similar test techniques used to assess language performance. In the light of this literature, some of the *objective test techniques* can be listed as in the following: (1) Multiple-choice, (2) Short-answer, (3) True-false, (4) Matching, (5) Completion, (6) Cloze test, (7) C-test, (8) Cloze elide tests, (9) Ordering tasks (Rearrangement), (10) Error correction, (11) Transformation, (12) Combination and Addition, (13) Word changing.

Among these objective test techniques, multiple-choice questions are commonly used by the test constructors. In fact, the SAMC questions are used more frequently than the MAMC ones. When a multiple-choice question is under discussion, a SAMC questions is usually meant. Thus, in the following part, SAMC questions and their disadvantages are described in detail.

3. SINGLE-ANSWER MULTIPLE-CHOICE QUESTIONS: DISADVANTAGES

Multiple-choice questions take many forms. Hughes (1989:59) points out that "there is a *stem* and a number of *options*, one of which is *correct*, the others being *distracters*". The teacher can give the question either through an incomplete sentence or through a full question. In fact, Hughes (1989) described not the multiple-choice

question in general, but the SAMC question specifically. Following example demonstrates the things that make a single-answer multiple-choice question.

Stem → Do not phone your father until he _____ you.

Options / Responses /← Alternatives	**a)** phoned	**→ Distracters**
	b) will phone	
	c) is phoning	
	d) phones	**→ Answer / Correct Option / Key**

(Adapted from Heaton, 1988)

Besides grammar, the teacher can test students' knowledge of pronunciation by means of single-answer multiple-choice test items. In his article on *Testing Pronunciation,* Bobda (1993) gives an example as follows:

In each of the following sets of words, three words have the same sound and one does not. Circle the one that does not have the same sound with the others.

1. a) dull   b) bull   c) wool   d) pull

2. a) poor   b) pour   c) sure   d) tour

As it is clear from the examples, SAMC questions can test some other language skills and areas. SAMC questions can be used to test listening, vocabulary, grammar, reading and pronunciation, but they will not be useful for testing all language skills. Madsen's (1983:38) statement makes this fact clearer. To him, "while multiple-choice tests can be used successfully in testing grammar, they do not seem to work as well in testing conversational ability".

In addition to not being able to test all language skills, these tests have some other disadvantages, too. Firstly, it takes a long time to write effective multiple-choice test items (Alderson, 2000; Heaton, 1988; Hughes, 1989; Madsen, 1983; Salkind, 2006). In addition, it is easy for students to cheat. Cheating is facilitated, as there are only four or five options. Students can communicate with their friends non-verbally and give the correct answer using their body language (Hughes, 1989; Madsen, 1983). Furthermore, Hughes (1989) provides three more disadvantages of these tests as (1) guessing may have a considerable but unknowable effect on test scores. (2) This technique restricts what can be tested as it is really difficult to find

enough distracters for the correct structure to be tested. Finally, (3) backwash may be harmful. In some cases, students are trained in guessing rather than in learning the language. Thus, multiple-choice tests have a harmful effect on learning and teaching.

Taking all these disadvantages of SAMC questions into consideration, a practical way of using and scoring MAMC questions is suggested in the following parts in detail.

4. MULTIPLE-ANSWER MULTIPLE-CHOICE QUESTIONS

Multiple-answer questions are rarely used while evaluating our students' success. Reviewing the literature, the author found out three different uses of multiple-answer questions. Firstly, MAMC questions are used in the questionnaires. For example, in a questionnaire asking people's hobbies, a person can choose more than one choice. However, in such uses, the constructor of the questionnaire never aims to evaluate students' success, but aims to get information about them. In the second use of MAMC questions; "there may be more than one correct answer but only one of them is the best of all the correct ones" (Salkind, 2006:143). Finally, in the last use, the tester provides four or five alternatives above the stem. Then, he/she asks the students which of the alternatives above are correct. Although there may be more than one correct answer in such questions, the student is supposed to choose *one option* which includes all the correct alternatives (e.g. Choices of Questions: a) I – II, b) II – III, c) I - III, etc.).

MAMC questions having more than one separate correct answer are hardly ever used in our classrooms. Very limited research has been carried out so far on MAMC questions (Duncan & Milton, 1978; Hsu & Moss & Khampalikit, 1984). Kurz (1999:15) also defines *multiple-answer format* as a test technique in which "examinees are instructed that any number of the options might be correct. Each item in this format is scored by giving the number of answers correctly marked minus the incorrectly marked options". In the light of Kurz's definition, the author developed a multiple-answer question format and a practical way of scoring it with which the

number of correct answers, the number of the distracters, the amount of the plus points for the correct answers and the amount of the minus points for the incorrect ones became much clearer. The advantages and disadvantages of this practical way of testing and scoring, and how it is to be scored are discussed in the following parts.

4.1 Using Multiple-answer Multiple-choice Questions: Some Advantages

The format of the MAMC question suggested in this paper includes a *stem* and *six options*, two of which are *correct*, and the others being the *distracters*. An example of a MAMC format can be seen in the following table:

Stem → If you've got a headache, take _____.

Options / Responses /← Alternatives	a) at home	→ *Four* Distracters
	b) in bed	
	c) to the doctor	
	d) to the dentist	
	e) some painkillers	→ *Two* Answers / Correct Options / Keys
	f) the day off	

Using MAMC questions can eliminate some of the disadvantages of SAMC questions. They are more advantageous in some aspects when compared to SAMC questions. Firstly, a tester can eliminate the decreased validity of SAMC tests due to guessing. While, in SAMC questions, the chance of guessing the correct answer is %33 (1/3), it is nearly %7 (1/15) in MAMC questions. The possibility of guessing two correct answers by chance in MAMC is almost impossible.

Furthermore, when *conventional number-right scoring method* is used, SAMC questions fail to credit partial knowledge. However, in a MAMC question suggested in this paper, the examinee can mark all the choices he/she wants. Hence, he/she can get some points for their partial knowledge (The scoring method and how students' partial knowledge is rewarded will be explained in the following section in detail).

In addition, it is easy for the students to cheat in exams including SAMC questions as there are only four or five options and a single correct answer. Students not only can communicate with their friends nonverbally to give the correct answer (Hughes, 1989; Madsen, 1983), but also can see the answer sheets of their friends sitting nearby. Though some teachers prepare different groups of exams each of which includes the same questions in different places to prevent cheating, it is usually easy to see the correct answer wherever the students sit and whatever group of exam they take as there is only one correct answer to be marked in their answer sheets. This is not the case in MAMC format which is suggested in this paper. As the students can mark as many choices as they want, this will cause great ambiguity in their answer sheet. Thus, no matter how clear a student sees his/her friend's answer sheet, he/she will not be able to detect which answer is correct or which one is wrong. As well as guessing the correct answers, cheating to find the correct answers is quite difficult in this format. Finally, in SAMC questions, it is difficult to find enough stems for the correct structure to be tested. However, by using MAMC, tester can ask two questions in one single stem. By using this format, it can be argued that the tester kills two birds with one stone considering the number of the stems used.

So far, many methods of scoring multiple-choice test items have been used to eliminate these disadvantages. Nearly all of these methods are used to score SAMC questions. These tests are traditionally scored using a *conventional number-right scoring method* in which items are scored with a value of '1' given to correct answers and a value of '0' given for incorrect, blank and omitted ones. Though this method is simple to use, it has the weaknesses discussed above. Therefore, an alternative way of scoring MAMC questions to eliminate the disadvantages of SAMC questions is suggested in the following.

4.2 Alternative Way of Scoring MAMC Questions: SAC Method

While scoring MAMC questions with six options two of which are correct, teachers are suggested to give a value of '+2' to each correct answer, a value of '-1' for each incorrect one and a value of '0' to each blank/unmarked alternative. In this

method, students can mark all the alternatives if he/she wants, but gets exactly the same score as if he/she did not answer the question. Students leaving all the alternatives unmarked (0) gets the same score when compared to students marking all the alternatives [+4 points (2 correct options) - 4 points (4 incorrect options) = 0]. As students are allowed to mark all the alternatives and the teachers are supposed to score all choices, this method can be named as *'Scoring All Choices'* (SAC) method. The following table exemplifies the way teachers should score students' answers:

Correct Options	Q.	A Examinee's Responses						A Examinee's Score for Each Question
A – C	1	A	B	C	D	E	F	4 points
B – D	2	A	B	C	D	E	F	3 points
E – F	3	A	B	C	D	E	F	2 points
D – E	4	A	B	C	D	E	F	1 point
A – F	5	A	B	C	D	E	F	0 point
B – D	6	A	B	C	D	E	F	2 points
C – F	7	A	B	C	D	E	F	1 point
A – E	8	A	B	C	D	E	F	0 point
C – B	9	A	B	C	D	E	F	-1 point
A – E	10	A	B	C	D	E	F	-2 points

As it can be clearly seen in the example above, MAMC questions both increase validity as it prevents guessing and is successful in crediting partial knowledge. For example, in *Question 2,* the examinee finds the first correct choice. However, he is in dilemma for the second one. Instead of not marking the second choice, the examinee can mark both the second correct answer and the strongest distracter. Therefore, he/she gets 1 point more for his/her partial knowledge. It is also clear in *Question 4* that the more alternatives the examinees guess to be true, the more points they lose. Thus, this method also prevents guessing while penalizing the incorrectly marked options.

4.3 Limitations of Multiple-answer Multiple-choice Questions

In addition to their advantages, such questions may have some drawbacks in language teaching. Firstly, finding two correct answers is much more difficult than finding two more distracters for the testers. As it is difficult to find two correct answers in one single stem, it is impossible to test all language skills and areas with MAMC questions. This method can only be an alternative way of testing some specific areas of language. Secondly, like SAMC questions, preparing effective MAMC questions can be really time-consuming. Finally, not scoring the responses carefully can mislead the results of the exam which will decrease its validity.

5. CONCLUSION AND IMPLICATIONS FOR ELT

Like many other objective test techniques (Short-answer, True-false, Matching, Cloze test, Ordering tasks, etc.), MAMC questions cannot test all the areas in English Language Teaching. However, when carefully prepared and organized, they can be a good alternative for testing some language areas being taught.

While testing grammar, the knowledge of the time expressions (two weeks ago, last week, etc.), the use of conjunctions (however, but, nevertheless, etc.), quantifiers (a lot of, plenty of, some, etc.), countable and uncountable nouns, tenses, modals (can, may, might, etc.), and passives can be tested using this new method of testing and scoring in our classroom. In testing vocabulary, this question type is most applicable to testing collocations, phrasal verbs, synonyms, and so on. The testers can use this form of questions in testing not only grammar and vocabulary, but also reading and pronunciation. Homophones and minimal pairs can be tested best using this alternative way of testing.

All in all, this alternative way of testing and scoring MAMC questions is open to discussion. The author does not claim a best method while providing this alternative. There is nothing best, but something better which will help the further developments in testing and scoring the examinees' success.

REFERENCES

Alderson, J. C. (2000) *Assessing Reading.* Cambridge: Cambridge University Press.

Bobda, A. S. (1993) Testing Pronunciation. *English Teaching Forum Online. 31*(3), 18-22. Available at:
http://exchanges.state.gov/forum/vols/vol31/no3/p18.htm

Duncan, G. T., & Milton, E. D. (1978): Multiple-choice test items: Responding and scoring through Bayes and minimax strategies. *Psychometrika, 43,* 43-57.

Genesee, F. & Upshur, J. A. (1996) *Classroom-based Evaluation in Second Language Education.* Cambridge: Cambridge University Press.

Heaton, J. B. (1988) *Writing English Language Tests.* London: Longman.

Hsu, T., Moss, P. A. & Khampalikit, C. (1984). The merits of multiple-answer items as evaluated by six scoring formulas. *Journal of Experimental Education, 36*, 152-158

Hughes, A. (1989) *Testing for Language Teachers.* Cambridge: Cambridge University Press.

Kurz, T.B. (1999). "A review of scoring algorithms for multiple-choice tests". Paper presented at the annual meeting of the Southwest Education Research Association. San Antonio, TX. (ED428076) Available at:
http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019 b/80/17/66/00.pdf

Madsen, S. H. (1983) *Techniques in Testing.* Oxford: Oxford University Press.

Salkind, N. J. (2006) *Tests and Measurement for People Who (Think They) Hate Tests and Measurement.* Thousand Oaks, California: Sage Publications.

Weir, C. J. (1990) *Communicative Language Testing.* New York: Prentice Hall.