

Mirjana M. Vučković*

University of Belgrade
Faculty of Philology
Serbia

Milica Prvulović**

University of Belgrade
Faculty of Philology
Serbia

THE AGONY OF CHOICE – EXPLORING STUDENTS' ANSWERS IN MULTIPLE CHOICE TESTS

Abstract

Multiple choice tests, in which students have to choose the correct answer from a limited set of options, belong to the category of selected-response assessments. They are considered to be relatively objective, fast and easy to mark. Most teachers believe that students should not change their first choice, although they are often in two (or more) minds about which answer is correct. In this research we tried to see whether this widely accepted belief was true and whether there was a connection between the student's level of knowledge and how prone he/she would be to changing his/her mind. We analyzed 342 tests taken in October 2014 by the students at the Faculty of Philology, University of Belgrade. The results showed that students with higher test scores changed their minds as frequently as those with lower test scores. On the other hand, the former were more likely to benefit from changing their minds than the latter.

Key words: *multiple choice tests, selected response assessment, testing*

* E-mail address: mirjana.m.vuckovic@gmail.com

** E-mail address: prvulovic.milica@yahoo.com

1. Introduction

According to Weitzman (1970: 71), “A multiple-choice item is a question presented with a number of possible answers or an incomplete statement accompanied by a number of possible completions, only one of which is correct. A person responding to the item ordinarily must make a single choice among the alternatives offered.”

Since they are given a choice of possible answers, students are often in two (or more) minds about which one is correct. Most teachers believe that the first chosen answer is usually correct, and students are often advised not to change their first choice. In this research we will try to see whether this widely accepted belief is true in the case of the students who chose English as their subsidiary subject at the Faculty of Philology, University of Belgrade. We analyzed 342 tests, which consisted of 30 multiple choice vocabulary and grammar tasks, taken in October 2014. Our goal was to investigate whether there were gains to be made by changing answers. We deem this to be important since learning takes place even while students are being tested.

2. History and development of MCTs

It is believed that the first multiple choice tests (MCTs) were used by an American psychologist Edward Thorndike (1874-1949) who is considered to be the father of modern educational psychology (https://en.wikipedia.org/wiki/Multiple_choice). His work led to the creation of tests which were introduced in 1917 and used to evaluate both intellectual and emotional characteristics of World War I recruits. Literate recruits took Army Alpha tests (Picture 1) which consisted of a number of questions with several answers to choose from, while the illiterate ones were tested with Army Beta tests (Picture 2) which consisted of pictures. These tests are considered to have been the first systematic MCTs.

Picture 1: Army Alpha

TEST 8

Notice the sample sentence:
 People hear with the eyes ears nose mouth

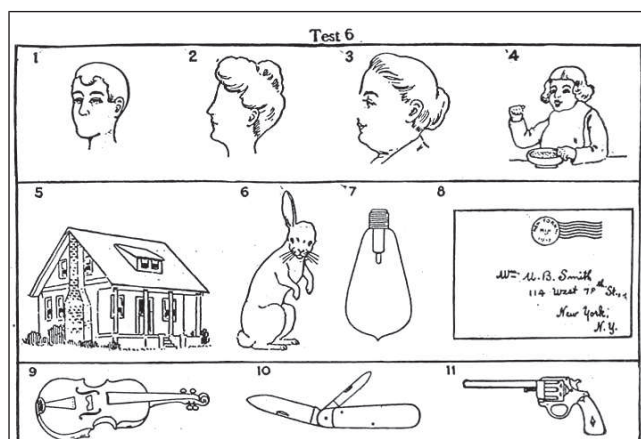
The correct word is ears, because it makes the truest sentence.

In each of the sentences below you have four choices for the last word. Only one of them is correct. In each sentence draw a line under the one of these four words which makes the truest sentence. If you can not be sure, guess. The two samples are already marked as they should be.

SAMPLES { People hear with the eyes ears nose mouth
 France is in Europe Asia Africa Australia

1	The apple grows on a <u>shrub</u> vine bush tree.....	1
2	Five hundred is played with <u>rackets</u> pins cards dice.....	2
3	The Percheron is a kind of <u>goat</u> horse cow sheep.....	3
4	The most prominent industry of Gloucester is <u>fishing</u> packing <u>brewing</u> automobiles..	4
5	Sapphires are usually blue <u>red</u> green yellow.....	5
6	The Rhode Island Red is a kind of horse granite cattle <u>fowl</u>	6
7	Christie Mathewson is famous as a writer artist baseball player <u>comedian</u>	7
8	Revolvers are made by Swift & Co. Smith & Wesson <u>W. L. Douglas</u> B. T. Babbitt.	8
9	Carrie Nation is known as a singer temperance agitator <u>suffragist</u> nurse.....	9
10	"There's a reason" is an "ad" for a <u>drink</u> revolver flour cleanser.....	10
11	Artichoke is a kind of <u>hay</u> corn <u>vegetable</u> fodder.....	11

Picture 2: Army Beta¹



¹ Both pictures were retrieved from (July 15, 2015.): https://www.google.rs/search?q=army+alpha+test&rlz=1C1AVNC_enRS573RS573&tbm=isch&tbo=u&source=univ&sa=X&ved=0ahUKEwjLt5fltOPJAhWrnnIKHcnWA2oQsAQIJQ&biw=1366&bih=643#imgrc=_

The popularity of MCTs grew during the twentieth century, and it especially increased since the sixties (Jarnagin & Harris 1977: 930), when teachers started using them in assessing the progress their students had made in various fields because they “permit a greater breadth of topics [...] and reduce the grading workload” (Jarnagin & Harris 1977: 930). These tests are believed to be “useful for measuring a fairly wide variety of different kinds of precise learning points” (Brown 1998: 659).

3. MCTs in learning languages – advantages and disadvantages

MCTs are especially popular in the field of language teaching. According to Brown (1998: 653), MCTs belong to the category of selected response assessments, together with true-false and matching exercises. In these three types of tests, test-takers have to search through a limited set of options and choose the correct answer (Brown 1998: 659). Thornbury (2010: 228) groups MCTs together with gap filling, and believes that their popularity lies in practicability. MCTs are considered as more appropriate for measuring receptive skills like listening and reading but should be avoided for productive skills like writing and speaking (Brown 1998: 658-659).

Another advantage of MCTs is that they are fairly reliable. Although, according to Edge and Garton, “no test is or can be perfectly valid or perfectly reliable” (2012: 162), a test is considered to be reliable if it produces similar results under similar circumstances, that is to say “[...] if two people of the same ability do the test, or if the same person does it twice, they should score the same” (Edge & Garton 2012: 162). Apart from these tests being relatively objective, their scoring is also fast and easy (Valette 1967: 88). We certainly know from personal experience that even hundreds of tests can be easily marked with a well-made key, since teachers do not waste time trying to decipher their students’ handwritings.

On the other hand, MCTs have certain disadvantages. These tests are very difficult to write, as all language teachers know. The distractors offered must be “plausible alternatives” (Smith 1982: 212), unless we want the item to be too easy to solve. Also, one needs to make sure that there is only one correct answer, even in various contexts. Furthermore, the correct answer should not be noticeably longer or shorter than the distractors or stand out in any other way (Edge & Garton 2012: 163). The “nature and difficulty of an item may be altered dramatically by changing the set of

distractors for the item” (Smith 1982: 212), so it can be quite demanding to come up with suitable distractors.

One of the objections against MCTs is that students can sometimes guess the correct answers and achieve higher scores (Valette 1967: 87). Jarnagin and Harris also believe that “a correct answer might be attributable to chance” (1977: 930). Another objection could be that cheating in these tests might be easier – this we know from personal experience.

Although there are some test-takers who find this format confusing, the others prefer it and find it easier since it contains the correct answer as well as the clues for choosing it (Smith 1982: 210). Our students usually prefer taking MCTs. Apart from having the advantage of recognizing the right answer when they see it, they probably also like the fact that no spelling or pronunciation mistakes can be made. When it comes to changing one’s answer, teachers usually believe that the first chosen answer is correct, and they often advise students not to change their answers. According to previous research (Reiling & Ryland 1972: 67), most students actually benefit from changing their initial answers. Their research showed that “final grade, sex of the student [...] had no significant impact on gains from changing responses” (1972: 67). We will try to find out how all this applies in the case of students who chose English as a subsidiary subject at the Faculty of Philology, University of Belgrade, but we will focus on the differences in successful answer-changing according to their levels of knowledge. Since MCTs are used in various fields, we believe that teachers of other subjects like mathematics, history, biology etc. could also benefit from the findings of this study.

4. Participants and test

The research was conducted with 342 students of various languages at the Faculty of Philology, University of Belgrade. All of them were students of 1st, 2nd or 3rd year, aged between 18 and 25. The participants had chosen to attend the first year of English as a subsidiary course and were given a short MCT to determine their level of knowledge. The answers they gave in this test were analysed in the research.

Picture 3: Test samples

7. Your colleague _____ call you until they _____ the meeting.
 A) won't / will finish B) going to / finish
 C) won't / finished ~~D) is not going to / have finished~~

8. Please make her boyfriend _____ her _____. My heart will break if she leaves.
 A) ask / stay ~~B) to ask / to stay~~ C) asking / staying ~~D) ask / to stay~~

9. Jake was exhausted yesterday and _____ down immediately.
 A) lied B) lain ~~C) lay~~ ~~D) laid~~

10. My _____ make very loud parties. They've decided to call the police.
~~A) friends' neighbours~~ B) neighbours of friends
 C) neighbours of friend's D) friend's neighbours

11. "Sheila _____ you that she can't see you because she is engaged."
~~A) regrets to tell~~ B) remembers to tell
~~C) reminds of telling~~ ~~D) regrets telling~~

12. Sandy _____ us that she had _____ this before.
 A) told / told B) said / told ~~C) told / said~~ D) said / said

20. _____ the 16th century tobacco was brought _____ Europe.
~~A) In / to~~ B) At / in ~~C) In / in~~ D) At / to

The test, which lasted for 35 minutes, consisted of 30 grammar and vocabulary items. All of them were multiple choice items with 4 options to choose from. Only one option in each item was correct. Before the test started the students were told by the invigilators that they could change their answers as many times as they wanted, as long as it was clear which answer represented their final choice. Picture 3 shows two samples of the described test and how the students marked their answers after changing their minds.

If we observe items 7 and 9, we will see what usually happens when students do MCTs. In item 7 the student chose answer D, and this was the correct answer. In item number 9 the student opted for answer D, although C was the correct answer. So, the student got one point for answering item 7 correctly, and lost one point for making a mistake with item number 9.

However, since our students were allowed to change their initial answers, we found a wide range of different examples while we examined the tests. If we look at item number 20, we will see that the student doing this test chose answer C, and then changed his or her mind, opting for answer A, which in the end proved to be the correct answer. And so, in this case, the student benefited from changing his or her mind. Next, in item

number 8 we have a situation where the student first chose answer D and then B, but actually D, the first choice, was the correct one, so the student did not benefit from changing his or her mind. Finally, if we look at item number 11, we see that the student changed his or her mind, but both times the answers he or she chose were wrong, i.e. in this particular case the process of changing his or her mind was neither beneficial nor harmful to the final test score.

Having observed the existence of these different outcomes which occurred when the students were allowed to change their answers, we assumed that there must be something causing some of the students to change their minds, and benefit from these changes. We thought this was due to the students' overall knowledge of English and their performance in this particular MCT. We therefore divided them into two groups according to their test scores: group A consisted of students who had higher test scores (30 to 26 points, 30 being the maximum number of points) and group B consisted of students who had slightly lower test scores (21 to 25 points). There were 151 students in group A and 191 in group B, which makes a total of 342 students.

5. Hypotheses

Before analysing the data from the tests, we had made the following assumptions:

1. The students in group A would change their minds about the answer fewer times than the students in group B.
2. The students in group A would be right to change their minds (i.e. their second choice would be the correct one, and they should therefore be encouraged to think twice and question their initial answers), whereas the students in group B should stick to their initial answers.
3. When the students in group B changed their minds, they would often choose an incorrect answer, alongside the incorrect choice they had made the first time, whereas this would rarely or almost never happen to the students in group A.

With our small exploratory study we also aimed to answer an important question: What advice should teachers give to their students who are taking a MCT?

6. Results and discussion

To compare how changing their answers correlated with their levels of knowledge, we divided the participants into two groups (A and B) according to their scores on the test. From Table 1 we can see that 151 (around 44%) students had high test scores – 26 to 30 points – and these were included in group A. Group B consisted of 191 students (around 56%) with lower test scores – 21 to 25 points. Table 1 represents the number of students in correlation to whether they changed their initial answers at least once or not at all.

Table 1: Number of students

	Group A	Group B	total
Students who changed their minds at least once	75	110	185
Students who did not change their minds	76	81	157
Total number of students	151	191	342

As can be seen from Table 1, almost exactly half of the students from group A (75 of them) changed their initial answers at least once, while the other half (76) did not change any of their initial answers. When we observe the numbers for the students belonging to group B, we can notice that 110 of them (which is around 58%) changed their minds once or more, while 81 (around 42%) did not. This means that the first hypothesis we postulated (that students with higher test scores would change their minds less frequently) was not correct. The difference between the two groups was only 8%, i.e. both groups of students, those with better and not so good knowledge of English, had an equal tendency towards questioning their answers in MCTs and changing them. We would like to point out that these numbers do not say how many times a particular student changed his or her answer in a particular test. We have the data for students who changed their minds at least once, and maybe many more times in a single test, and those who did not change their minds at all. This is because we were trying to investigate whether the students in group A or group B had a general tendency towards questioning their answers in MCTs and changing them.

Table 2 represents the numbers of actual changes made by the students as well as the outcome of these changes. In the 342 tests that were analysed there were 346 changes. 124 (around 36%) were made by group A, and 222 (around 64%) by group B. This means that, although both groups of students were inclined to change their answers in MCTs, students in group B were more likely to make several changes in a single test.

Table 2: Number of changes

	Group A	Group B	total
Changing their mind led to a correct answer	17	128	145
Something else was correct (a 3 rd option)	12	48	60
Changing their minds – TOTAL	124	222	346

First, we will analyse the results for group A. From Table 2 we see that the 75 students from group A who changed their minds (see Table 1) made a total of 124 changes in their tests. Some of them made just one change, some of them more. The student from this group who changed most of his or her answers in a single test did so 6 times. Out of the 124 changes, 95 proved to be good choices, i.e. the student's second choice was the correct answer to the test question. In only 17 cases the students from group A made the correct choice the first time they answered the question. Even fewer (only 12) made a completely wrong choice altogether, i.e. they chose one answer which was incorrect, then changed their minds, again choosing the incorrect answer. If we view this through percentages, we get the result that almost 77% of the time the students in group A were correct to change their first answers. This proves the first part of our second hypothesis: students who achieve better test scores can benefit from questioning their initial answers in MCTs and changing them if they feel their first choice was incorrect.

Next, we will analyse the results for the group of students with lower test scores (group B). The 110 students from group B who changed their minds once or more (see Table 1) made a total of 222 changes, which is almost twice as many times as the other group (one student from this group made as many as 8 changes in a single test). The students were right to change their answers in only 46 cases, which is around 21%. Many of the students belonging to group B chose the correct answer the first time,

and then changed it (128 changes of this kind were made, which is around 58%). This proves the second part of our second hypothesis: students with lower test scores should not be encouraged to change their minds in MCTs, but stick to their initial answers. There were also many of them in this group who made a completely wrong choice altogether, choosing the wrong answer both the first and the second time (48 mistakes of this type were made, which is 4 times more than in group A). This proves our third assumption: students from group B made wrong choices both before and after changing their minds.

7. Conclusion

In this paper we have discussed and roughly illustrated how students' general knowledge of English and their test scores influence their inclination towards second-guessing their answers in MCTs and changing them, as well as the consequences of those changes. We have found all of our hypotheses, except the first one, to be true. Both the students with better knowledge of English and those with a lower level of knowledge are prone to changing their minds in MCTs. Next, students with higher test scores should be encouraged to question their initial choices in MCTs, as they are likely to benefit from changing their minds. Finally, when students with lower test scores change their minds, they often have the wrong answer to begin with, so it makes no difference whether they change their minds or not and should therefore be neither encouraged nor discouraged to change their minds.

Our study sought to answer one main research question: What advice should teachers give to their students who are doing a MCT? According to Clark (1970: 21), "intuitive problem-solving is an aspect of education neglected almost everywhere except in multiple choice tests [...] it needs to be understood that [...] the quality of intuitive insight is directly related to the amount of knowledge brought to the problem." The results we obtained indicate that teachers should advise the students with better test scores and better knowledge of the English language to analyse and second-guess their initial answers when they are not sure if they are right. On the other hand, the students who tend to get lower test scores and who are not very good at English should be told to use their intuition and stick to the answer they thought was correct the first time around. We admit

that such an approach is highly controversial, and we do not recommend it unless some other possible explanations are explored. For example, there is some evidence that significant differences in test taking exist between boys and girls. Gurian et al. (2001) claim that MCTs are easier for boys than girls, that boys are more likely to take a chance and guess the answer, whereas girls prefer to be certain about their answers. Maybe there are gender differences in students' tendencies to change their answers in MCTs and the outcome of those changes. We believe that further research in this field would prove beneficial to helping students achieve higher scores in their tests as well as learn English more efficiently.

References

- Brown, J. & Hudson T. (1998). The Alternatives in Language Assessment. Vol. 32, No. 4 Teachers of English to Speakers of Other Languages Inc. (TESOL). 653-675.
- Edge, J. & Garton, S. (2012). *From Experience to Knowledge in ELT*. Oxford University Press.
- Gurian, M., Henley, P. & Trueman, T. (2001). *Boys and Girls Learn Differently!* San Francisco, CA: Jossey-Bass. (June 5, 2015) <http://www.slideshare.net/vikramjeet925/boys-and-girls-learn-differently-a-guide-for-teachers-and-parents>
- Jarnagin, B. & Harris J. (1977). Teaching with Multiple Choice Questions. *The Accounting Review*, Vol. 52, No. 4. American Accounting Association. 930-934.
- Nettleton, K. F. (2008). Fair Game: Gender Differences in Educational Games. In Miller, C. (Ed.) *Games: Purpose and Potential in Education*. New York: Springer. 55-72.
- Reiling, E. & Ryland T. (1972). A new approach to the Problem of Changing Initial Responses to Multiple Choice Questions. *Journal of Educational Measurement*, Vol. 9, No. 1. National Council on Measurement in Education. 67-70.
- Smith, Jeffrey K. (1982). Converging on Correct Answers: A Peculiarity of Multiple Choice Items. *Journal of Educational Measurement*, Vol. 19, No 3. 211-220.
- Thornbury, S. (2010). *An A-Z of ELT: A dictionary of terms and concepts used in English Language Teaching*. Macmillan Books for Teachers.

Valette, R. (1967). Improving Multiple-Choice Grammar Tests in German. *The German Quarterly*, Vol. 40, No. 1. Wiley on behalf of American Association of Teachers of German. 87-91.

Weitzman, R. A. (1970). Ideal Multiple-Choice Items. *Journal of the American Statistical Association*, Vol. 65, No 329. Taylor and Francis Ltd on behalf of American Statistical Association. 71-89.