Partial Knowledge Reduces Performance in Multiple Choice Tests

Zekeriya Eser
Department of Accounting, Finance, and Information Systems
School of Business
College of Business and Technology
Eastern Kentucky University
Richmond, KY 40475, USA.

Abstract

We find that adding a weaker distractor to a two-option multiple choice item decreases student performance through two channels. First, the new option increases students' confusion about the right answer. Conditional on being confident about the answer, there is not much statistically significant decrease in student performance. The second channel operates when students are not fully confident of their answer. This is when student performance may decline due to misleading partial knowledge.

Keywords: Partial knowledge, confidence-based testing, low cognition, guessing, multiple-choice

Introduction

While the results vary in the literature (Tversky, 1964; Rodriguez, 2004; Rodriguez et al, 2014) at least up to a point, increasing the number of alternatives in a multiple-choice test tends to decrease student performance. If students answer a portion of questions by random guesses, increasing the number of options in a multiple choice item will decrease their chances of getting lucky, hence their performance. An alternative explanation is that adding one more distractor creates a new attractor for students with lower cognitive abilities.

Since the results in the literature are based on traditional multiple choice items, it is impossible to discriminate between competing explanations based on the results of such traditional tests. In any case it is impossible to know if students were guessing randomly or based on partial knowledge. Building on our earlier paper that employed Confidence Based Testing (CBT) (Colbert et al, 2012), we offer a novel experimental design where the role of pure uncertainty can be (somewhat) filtered out. In addition to traditional multiple choice items, students are asked the same questions in the CBT (we will refer to CBT questions as Fuzzy from now on) format where not only they are asked to provide a correct answer but also their confidence in that correct answer. There are extra rewards for being fully confident and giving the correct answer, but there are also severe penalties in the form of negative scores for being fully confident and providing an incorrect answer. Students also have the option of declaring ignorance for which they get a small reward.

We find that adding a weaker distraction to a two-option multiple choice item decreases student performance through two channels. First, the new option increases students' confusion about the right answer. Students are no longer as certain of the answer they would have picked when there was only two-options. The decrease is about 11 percentage points. Conditional on being confident about the answer, there is not much statistically significant decrease in student performance. The second channel operates when students are not fully confident of their answer. This is when student performance may decline due to misleading partial knowledge. Even after adjusting for increased difficulty of guessing we find that performance decline is over 16 points in a 100 point exam.

Experiment Design

Our experiment starts with a multiple-choice question selected from the test banks provided by the textbook publishers. These questions have five options. Based on the instructors' judgment we identify the correct answer, and rank the incorrect answers by the strength of their distracting power.

Two of the options that are deemed to be the weakest distractors are eliminated. Using the correct option, the best distractor, and the second-best distractor, we form four questions:

- A two-option traditional question comprised of the correct answer and the best distractor
- A three-option traditional question comprised of all three surviving options
- A two-option fuzzy question comprised of the correct answer and the best distractor
- A three-option fuzzy question comprised of all three surviving options

Students are randomly assigned one of these four questions. We repeat this procedure for all of the items selected for the exam.

Basic Results

We first present the results from the traditional questions (A and B questions).

	Success Ratio in Traditional Multiple-Choice Questions		
	Raw	Guess-Adjusted	N
2-option	74.58%	49.15%	354
3-option	60.41%	40.61%	394

Using the traditional multiple-choice format, students do significantly worse—the t-ratio for the difference is over 4.19—in the 3-option items compared to 2-option items. The difference in the raw performance is about 14 points in a 100 point exam. However, some of that performance decline may be attributable to pure guessing when student does not have any idea about the correct answer. Even after adjusting for the lower probability of successful guesses in the three-option items, the difference is still over 8.5 points in a 100 point exam—with a t-ratio of 2.35 for the difference.

We obtain the guess-adjusted success ratios as follows. Let x represent the percent of questions student answered believing that it was the correct answer, let y represent the percent of questions student actually got right, and let n represent the number of options present in the item. Now, assume that for (100-x) percent of the questions students chose the answer by pure random guessing. Then, we can solve for x using

(1) $x + (100 - x) /n = y$, which implies $x = (ny - 100) / (n-1)$

When students are asked the multiple-choice items in the Fuzzy format we obtain information in multiple dimensions. We first examine the percentage of students who were fully certain of their answer, as well as performance conditional on being fully certain. The following table contains these results:

	Success Ratio in Multiple-Choice Items for Fuzzy Questions			
	Raw	Guess-Adjusted	Fully Certain	N
2-option	80.00%	60.00%	69.57%	240
3-option	75.64%	63.46%	58.50%	234

The first result to note in this table is that adding a weaker distractor to a two-option item decreases students' confidence in their answer. The difference is over 11 percentage points and is statistically significant with a t-ratio of 3.16. Conditional on students being fully confident in their answer, the reduction in performance for the three-option questions is much smaller; less than 4 points in a 100 point exam, and is not statistically significant with a t-ratio of 1.14. When student performance, conditional on being fully confident, is guess adjusted, students actually perform better in a three-option item by about 3.5 points even though the difference is not statistically significant with a t-ratio of 0.77.

We also impute the student performance in the questions where they were not fully certain of their answer. To see how we do this consider the performance in the two-option case. Students are fully confident of their answer for 69.57% of the questions. For these questions, they give the correct answer 80% of the time. For the remaining 30.43% of the questions, they give the correct answer for y. Imputation is done by forcing the weighted success rate to be equal to the success rate for the traditional format, which was 74.58%. Hence

(2) 0.6957 x 0.8 + 0.3043 y = 0.7458, hence y = 62.19% (guess-adjusted: 24.38%)

Similarly, for the three-option items students are fully confident of their answer for 58.50% of the questions. For these questions, they give the correct answer 75.64% of the time. For the remaining 41.50% of the questions, they give the correct answer for y. Imputation is done by forcing the weighted success rate to be equal to the success rate for the traditional format, which was 60.41%. Hence

(3) 0.5850 x 0.7564 + 0.4150 y = 0.6041, hence y = 38.94% (guess-adjusted 8.41%)

Hence, when students are not certain of their answer, the degradation in their performance in three-option items is much higher: over 23 points before adjustment, and about 16 points even after adjusting for difficulty of guessing due to increased options.

Concluding Remarks

Our results show that adding even weak distractors result in reduced performance in multiple choice questions by first leading students to lower confidence in their answers, and then reducing performance further due to low cognitive skills. Further study is needed to where we can control degree of difficulty of questions as well as student skill levels to highlight the role of partial knowledge itself in degrading student performance.

References

Colbert, J., Eser, Z., Holbrook, M.E. (2012) "Confidence Based Marking: Implementation and Feedback Measures" Journal of Higher Education Theory and Practice, Vol. 12 (1), pp. 27-38.

Rodriguez, M.C. (2005) "Three Options Are Optimal for Multiple-Choice Items: A Meta-Analysis of 80 Years of Research" Educational Measurement: Issues and Practice, Vol. 24 (2), pp. 3-13.

Rodriguez, M.C., Kettler, R.C., Elliott, S.N. (2014) "Distractor Functioning in Modified Items for Test Accessibility" SAGE-Open, Vol. 4 (4).

Tversky, A. (1964) "On the Optimal Number of Alternatives at a Choice Point" Journal of Mathematical Psychology, Vol. 1 (2), pp. 386-391.