Winter 1991

The Myth of "Conventional Wisdom" on Changing Multiple-Choice Answers

Marshall A. Geiger
University of Richmond, mgeiger@richmond.edu

The Myth of "Conventional Wisdom" on Changing Multiple-Choice Answers

Marshall A. Geiger
Assistant Professor
University of Maine

ABSTRACT

Business students are often warned not to change multiple-choice answers once an original selection has been made. This "conventional wisdom," that the first answer selected usually is the correct answer, is in contrast with the conclusions of research in the education and psychology fields. This study extends these earlier studies by using students in accounting principles I and principles II classes, and by examining whether the type of question (numeric or non-numeric) affects answer-changing behavior. On average, for every point lost roughly three points were gained by changing answers for both groups. Additionally, gender was found not to be a factor on the net point gain/loss of the student. Question type did not influence the overall tendency to change answers for the principles I group; however, the principles II students changed fewer numerical question answers than non-numerical answers. Also, some tendency was shown for males to change more answers than females when type of question and direction of change were analyzed. Overall, the findings clearly evidence the benefit of changing multiple-choice answers if a student believes his/her original selection to be incorrect.

Classical test theory suggests that the primary determinant of student's response to items on a cognitive test is knowledge or aptitude, and that these attributes are reflected in final test scores. Research over the past several decades has identified a host of other factors, including item bias, test anxiety and testwiseness that can also effect exam performance. An additional characteristic that has been shown to effect exam performance is a student's answer-switching behavior on multiple-choice questions.

Although several studies have demonstrated the positive effect of changing answers, students often receive the advice from business educators not to change test answers once an initial answer selection has been made. This "conventional wisdom", that the first answer selected is usually the best answer, encourages students to maintain their first answer selection. Even though students are aware of the "don't change" warning, some students invariably continue to change original answer selections on multiple-choice exams. This research is intended to assess whether the "don't change" warning can be empirically supported using two groups of introductory accounting students. Additionally, it also analyzes whether the type of multiple-choice question affects answer-changing behavior, as

well as examines behavior differences due to the gender of the student.

PRIOR STUDIES

The phenomena of changing answers on objective examinations has been studied sporadically for the past 60 years. The first empirical study by Mathews (1929) examined college students in educational psychology courses and found that over 53 percent of the answers changed on multiple-choice questions were from a wrong answer to the right answer (WR), approximately 22 percent were from the right answer to a wrong answer (RW) and the remaining changes were from wrong answers to other wrong answers (WW). These findings led Mathews (1929) to conclude:

> Students should be informed that it pays in terms of score to check over all questionable items carefully in true-false and multiple-choice types of tests rather than to trust first impressions. They may expect to raise their scores at least twice as often as they lower them by changing their first responses when later judgment seems to justify it (p. 286).

The basic findings of the Mathews (1929) study, that roughly two to three points are gained for every point lost, have been upheld by several later researchers.[1] Additionally, subsequent researchers have examined several individual characteristics and potential causes of answer-switching behavior. Reile and Briggs (1952) found that females changed answers more often, but overall gained fewer points than their male peers. Bath (1972), however, found that females gained more points than males, and several studies have found no gender relation to switching behavior or net gain (Copeland, 1972; Mueller and Shwedel, 1975: Reiling and Taylor, 1972). Vidler and Hansen (1980) found that changes were more likely to be made on difficult rather than easy items. Jacobs (1972), however, concluded students changed more answers on low to moderate difficulty items. Jacobs (1972) also examined student's perceptions of changing answers and found that most believed they lost points by changing answers. When examining whether these aggregate perceptions were justified, he found that all groups gained points, even the group that believed they lost points. Additionally, Foote and Belinsky (1972) found that after the college students were aware of the potential benefit to changing answers, they still did not alter their answer-changing behavior (i.e. change any more answers) when taking multiple-choice examinations.

Unfortunately, most business educators today are generally unaware of these empirical findings and still encourage students not to change answers on multiple-choice questions once an original answer is selected. However, all of the prior studies have utilized non-business students. Prior research has been performed on elementary school students or college students in psychology, education or educational psychology courses. No investigation has been performed on college students outside of these closely related fields to assure that the results are generalizable to students in other disciplines. Confirmation of these overall results should come from other groups of students in different courses before they should be interpreted and applied generally.

This research is an attempt to partially validate and extend these early results by using students of introductory accounting. Not only are these students of another discipline, they are students of a discipline that at times poses fundamentally different questions. Accounting students are expected to have an aptitude with numerical, as well as non-numerical issues and problems. For the most part, the disciplines studied in the earlier research do not have a significant numerical component. The results obtained by the earlier studies, then, may have been driven by the type of questions comprising the examinations (i.e. non-numerical). Accordingly, this research has also separately analyzed answer-changing behavior on both numeric and non-numeric multiple-choice test questions, as well as examined for gender differences in answer-changing behavior.

METHODS

Sample

Two independent groups of introductory accounting students were used in this study. The first group consisted of 120 students (73 males and 47 females) taking an accounting principles I course. The second group consisted of 124 students (84 males and 40 females) taking an accounting principles II course. The students were predominantly (85% and 91%, respectively) business administration majors with concentrations in accounting, finance, management, marketing and management information systems. The introductory 3-credit courses were required for all business administration majors.

Items

Multiple-choice questions given on four in-class examinations were used for both groups. All examinations consisted of a multiple-choice section and a "problems" section. In total, 96 multiple-choice questions were given to the principles I students and 143 questions were given to the principles II students during the semester.

The multiple-choice questions were then identified as being either numeric (N); that is, questions involving a numerical calculation to arrive at the answer, or non-numeric (NN); that is, questions asking about concepts, classifications, definitions, etc. For the principles I students there were 35 numeric and 61 non-numeric multiple-choice questions throughout the semester, while the principles II students answered 54 numeric and 89 non-numeric questions. Due to the separate "problems" section given on the exams, more non-numeric (NN) than numeric (N) questions were given as multiple-choice items. The number of multiple-choice questions included in this study, then, was 11,520 (4,200 N and 7,320 NN) for the principles I students and 17,732 (6,696 N and 11,036 NN) for the principles II students for a total of 29,252 (10,896 N and 18,356 NN).

Procedure

All examinations were hand-graded and retained by the researcher. Students were asked to "circle the best answer" for the multiple-choice questions directly on the test instrument itself. The examinations were then scrutinized under high illumination for erasure marks. Since the actual examinations were used, evidence of answer-changing was readily apparent. In cases of uncertainty, the questionable item was not considered a change.

Unlike several earlier studies, this research used the actual examination instrument to identify changes. This was believed to provide a more accurate record of changing behavior than using machine readable answer sheets. Students can mark the actual exam and then change answers prior to transferring their final selections to the answer sheets, thus providing the researcher with no evidence of actual answer-changing behavior on the final answer sheet. All identified changes were classified as WR, RW, or WW, as well as whether the change was on a numeric (N) or non-numeric (NN) question.

RESULTS

Overall Changes

Table 1 presents a summary of the aggregate changing behavior of the two groups of accounting students. Overall, the combined percentage of answers changed from wrong to right (WR) was almost 57 percent, with the remaining changes fairly evenly split between right to wrong (RW) and wrong to wrong (WW) changes. Both principles I and principles II students evidenced very similar overall patterns. These overall results are consistent with those obtained in other disciplines and supports the notion that upon additional thought, changing original answer selections may well be in the best interest of the student. Overall, only 21 percent of the time did the change on an individual question cause the student to lose points, yet 57 percent of the time students gained points from making the change.

Table 1

Overall Summary

	RW	WR	WW	Total Changes
Principles I				
Number of Changes	142	404	128	674
Percent of Total (N = 674)	21.1	59.9	19.0	100.0
Principles II				
Number of Changes	160	421	202	783
Percent of Total (N = 783)	20.4	53.8	25.8	100.0
Combined				
Number of Changes	302	825	330	1457
Percent of Total (N = 1457)	20.7	56.6	22.7	100.0

Individual Student Changes

While the overall analysis supports answer-changing behavior, it is more from an individual perspective that the dilemma of whether or not to change an answer takes on meaning. Therefore, answer-changing data was gathered on an aggregate basis as well as an individual basis to identify how many individual students actually gained or lost points for the semester due to their answer-changing behavior. Additionally, the relationship of gender to answer-changing behavior has been addressed. Table 2 presents the results of answer-changing behavior by gender and in total for both groups.

Table 2

Mean Changes by Gender*

	RW	WR	WW	Total Changes	Net Gain
Principles I					
Male	1.178	3.589	1.096	5.863	2.411
N = 73	(1.398)	(2.847)	(1.180)	(4.467)	(2.425)
Female	1.191	3.021	1.021	5.234	1.830
N = 47	(1.209)	(2.111)	(1.294)	(3.205)	(2.099)
Overall	1.183	3.367	1.067	5.617	2.183
N = 120	(1.322)	(2.589)	(1.221)	(4.017)	(2.312)
Principles II					
Male	1.345	3.595	1.607	6.548	2.250
N = 84	(1.322)	(2.863)	(1.715)	(4.818)	(2.643)
Female	1.175	2.975	1.675	5.825	1.800
N = 40	(1.174)	(2.190)	(1.685)	(3.856)	(2.221)
Overall	1.290	3.395	1.629	6.315	2.105
N = 124	(1.274)	(2.671)	(1.698)	(4.527)	(2.514)

( ) *Standard Deviation*

* *Individual t-test results for each category indicate no difference due to gender of student at .10 level for both groups.*

Of the 120 principles I students, 115 (95.8%) changed at least one answer and of the 124 principles II students, 120 (96.8%) changed at least one answer for the semester. Table 2 indicates that the mean number of answer changes for the principles I students was 5.6 or approximately 6 percent, and for the principles II students it was 6.3 or approximately 4.4 percent of all answers. These tendencies to switch answers are higher than the average of 3.2 percent reported by Mueller and Wasser (1977) in their review and summary of the earlier answer-changing studies. However, one study by Reile and Briggs (1952) found that 6.2% of all answers were changed by the psychology students in their study. Thus, it appears that these groups of accounting students might be slightly more apt to change answers than their colleagues from other disciplines, but that any difference in the tendency to change answers appears minimal.

The mean net gain in points for the principles I students was 2.18 (or 2.27 percent), and for the principles II students was 2.11 (or 1.48 percent). These results of percentage gain are also consistent with earlier research. Although these accounting students may change slightly more answers, the residual effect on net gain for the semester was not materially altered. Tables 1 and 2 also indicate that, on average, for every point that any one student lost due to changing answers, they gained roughly three. The gain/loss ratio for the principles I and principles II students were 2.85/1 and 2.63/1, respectively. These results, too, are consistent with the two to three point gain/loss ratio found in prior studies.

Separate t-tests were performed to determine if there was an effect due to gender on each category of change, the total changes and the net gain for both groups. Each of the tests produced insignificant results at the .10 level. Accordingly, there is no evidence that males and females differ in their answer-changing behavior on these introductory accounting exams.

The standard deviation for number and type of changes in Table 2 are all fairly large, and in some cases are greater than the mean. These large variances give evidence of substantial differences in individual behavior. For example, the total number of changes made by any one principles I student ranged from zero to 26 (27% of all 96 questions). The maximum numbers of RW, WR and WW changes by any one principles I student were 7, 14, and 5, respectively, while several students changed no answers or only a few. Similar variances across students were also found in the principles II group. Thus, it is very difficult to predict a priori, on an individual basis, which students will gain or lose points and what the magnitude of net gain/loss will be. Table 3 presents the number of students, by sex, that actually gained and lost points during the semester due to changing their original multiple-choice answer selections.

Table 3

Summary of Individual Gains/Losses by Gender*

	Principles I			Principles II		
	Males N = 73	Females N = 47	Total N = 120	Males N = 84	Females N = 40	Total N = 124
Gainers	60	31	91	59	28	87
	82%	66%	76%	70%	70%	70%
Losers	5	5	10	11	5	16
	7%	11%	8%	13%	12%	13%

Table 3 - *Continued*

Summary of Individual Gains/Losses by Gender*

	Principles I			Principles II		
	Males N = 73	Females N = 47	Total N = 120	Males N = 84	Females N = 40	Total N = 124
No Change	8 11%	11 23%	19 16%	14 17%	7 18%	21 17%
Mean gain [a]	3.10 (2.014)	2.90 (1.729)	3.03 (1.924)	3.440 (2.223)	2.857 (1.715)	3.253 (2.081)
Mean loss [a]	2.00 (1.732)	1.00 (0)	1.50 (1.224)	1.273 (1.667)	1.600 (.548)	1.375 (.500)

( ) *Standard Deviation*

* *x^2 tests for dependence are not significant at the .10 level for either group.*

[a] *t-tests indicate no difference in sample means at the .10 level for either group.*

For the principles I students, sixty males (82%) and 31 females (66%) actually gained points over the course of the semester, while five males (7%) and five females (11%) had net losses of points. Similarly, for the principles II students, 59 males (70%) and 28 females (70%) gained points, while 11 males (13%) and 5 females (12%) lost points due to their answer changing behavior for the semester. Separate Chi-square tests of dependence on both groups indicate that net gain or loss of points for the semester is not dependent on the gender of the student ($p>.10$). Additionally, separate t-tests for the magnitude of gains and losses, by gender, also indicate that the size of the gain or loss is not dependent on the sex of the student. These results evidence that the gender of the introductory accounting student plays no significant overall role in the switching behavior of the student or the outcome of that behavior.

Type of Question

The accounting students might behave like the students in other disciplines overall, but still maintain different behaviors on different types of multiple-choice questions. Prior research utilized primarily non-numerical test questions for analysis. Each multiple-choice question, and change, in this research was identified as being numerical or non-numerical. Table 4 summarizes the switching behavior by question type for both groups.

Although the principles I students switched slightly more (as a percent) non-numeric than numeric questions, the differences across question type are not significant ($p>.10$) for any of the four categories. However, there does appear to be a difference in answer-switching behavior due to question type for the principles II group. These students switched significantly ($p<.01$) more non-numeric answers than numeric answers for each category of change. A closer examination of Table 4 indicates that the difference due to question type appears to stem from the lower number of numerical answer changes by this group. When compared to the results of the principles I students, the principles II students appear to switch the same proportion of non-numerical answers for each category of change (i.e., RW, WR, WW and Total), but that they are less apt to make numerical answer changes than their principles I counterparts.

Table 4

Mean Number of Switches by Question Type

Principles I (120 students)*

	RW			WR			WW			Total		
	# Changed	Mean	% Changed	# Changed	Mean	% Changed	# Changed	Mean	% Changed	# Changed	Mean	% Changed
Numeric N = 35	44	.367 (.634)	1.05	140	1.167 (1.140)	3.34	40	.333 (.639)	.95	224	1.867 (1.670)	5.33
Non-Numeric N = 61	98	.817 (1.085)	1.34	264	2.200 (2.069)	3.61	88	.733 (.959)	1.20	450	3.75 (3.099)	6.15
Overall N = 96	142	1.183 (1.322)	1.23	404	3.367 (2.589)	3.51	128	1.067 (1.221)	1.11	674	5.617 (4.017)	5.85

Principles II (124 students)**

	RW			WR			WW			Total		
	# Changed	Mean	% Changed	# Changed	Mean	% Changed	# Changed	Mean	% Changed	# Changed	Mean	% Changed
Numeric N = 54	25	.202 (.441)	.37	66	.532 (.950)	.98	29	.234 (.479)	.43	120	.967 (1.355)	1.79
Non-Numeric N = 89	135	1.088 (1.894)	1.22	355	2.863 (2.293)	3.22	173	1.395 (1.519)	1.56	663	5.347 (3.893)	6.01
Overall N = 143	160	1.290 (1.274)	.90	421	3.395 (2.671)	2.37	202	1.629 (1.698)	1.14	783	6.315 (4.527)	4.42

( ) *Standard Deviation*

* *Individual t-test results indicate no difference for each category due to question type at the .10 level for this group.*

** *Individual t-test results indicate significant differences for each category due to question type at the .01 level for this group.*

These results indicate that while there might not be an effect due to question type on students taking their first accounting course, the same may not be true for students in later courses. One potential reason for the lower number of numerical answer changes could be that students gain more confidence about their ability to perform numerical calculations as they get more exposure and practice performing them. This confidence may lead to fewer answer changes on numerical exam questions. These initial results on difference in answer changing behavior due to question type need to be validated by future studies in order to be generalized beyond this study.

Additionally, all categories of switches by question type (i.e., RWN, RWNN, etc.) in Table 4 were analyzed by gender for both groups. For the principles I students, two statistically significant results were obtained. First, males

were more likely than females to switch answers on numeric questions (p = .06). Males evidenced a mean of 2.08 numeric changes, while females maintained a mean of 1.53 changes. Second, males were more likely to switch answers on numeric questions from right to wrong than females (p = .03). Males demonstrated a mean of .466 RW numeric changes, while females maintained a mean of only .213 changes. This second statistical finding for this group appears to largely drive the first. All other categories of changes did not evidence a gender effect at the .10 level.

For the principles II students the only statistically significant difference due to gender for all categories was the wrong to right non-numerical category. Again males made more changes (mean of 3.15) than females (mean of 2.25). However, there was no statistical effect due to gender for the overall number of wrong to right changes (p = .19) or the number of non-numerical changes (p = .12). Hence, the combined evidence from the principles I and principles II groups indicates that males may be slightly more prone to switch answers than females when the most detailed level of analysis is used. However, any differences due to gender in answer changing were not manifest in significant differences in the total number of changes or in net point gain/loss for the semester. Further research is needed to replicate these initial results of gender and question type interaction before they can be generalized beyond this study.

Summary and Discussion

The overall results of this study are consistent with those obtained by earlier researchers in psychology, educational psychology, and education. Generally, prior research has found that students gain roughly two to three points for every point lost due to their answer-switching behavior. For this study, for every point lost due to switching, on average almost three points were gained for each group. The average net gain for the semester was approximately 2 percent for the principles I students and 1.5 percent for the principles II students. Ninety-six percent of all students changed at least one answer, and over 70% of the students in each group increased their scores over the semester due to their switching behavior. Additionally, no differences in mean net gain or magnitude of gains or losses were found to be due to the gender of the student.

This study also extended earlier research to analyze question type—numeric vs. non-numeric—and found no overall differences in answer-changing behavior for the principles I students. However, principles II students were found to change significantly fewer answers to numerical questions than non-numerical questions. Additionally, some evidence was presented that males are more likely than females to change answers when type of question and direction of change was analyzed.

The results of this study clearly indicate that accounting students should be encouraged to reconsider and evaluate their responses to multiple-choice test items. Millman, et. al (1965) and Sampson (1985) suggest that the tendency to evaluate and judiciously change one's response is a basic aspect of test-wiseness that can be enhanced by proper instruction. Students should, upon additional thought, be encouraged to change original answer selections if they believe the original selection is incorrect.

There still remain several issues for further research in this area. Along with replicating the present study, future research should attempt to assess individual traits of students and their effect on answer-changing behavior. Traits such as anxiety level or learning style have yet to be examined in conjunction with answer-changing behavior. The issue of whether students can accurately predict the outcome of their individual answer-changing behavior has yet to be addressed. Earlier research has shown that in the aggregate they can not accurately predict this outcome, but no studies have performed an individual analysis. Additionally, benefits of answer-changing on professional licensing examination questions offers another interesting extension in this area.

The areas of future research are directed toward ascertaining information on why the

answer-changing behavior of test takers works, and who it works best for, not on assessing whether it is generally beneficial. The results presented here, when coupled with those of researchers in other disciplines, present strong evidence that the "conventional wisdom" regarding answer-changing is not only incorrect, but is potentially detrimental to students who would otherwise rethink and change answers.

FOOTNOTES

[1]For a general review of this literature see Mueller and Wasser (1977).

REFERENCES

Ballance, Collin T. and W. Eugene Gentzel (1980) "Effect of Several Scoring Techniques on Answer-Changing By Students" *Psychological Reports*, Vol. 46, pp. 653-654.

Bath, John A. (1967) "Answer-Changing Behavior on Objective Examinations," *The Journal of Educational Research*, Vol. 61, No. 3, pp. 105-107.

Foote, Russell and Charles Belinky (1972) "It Pays to Switch? Consequences of Changing Answers on Multiple-Choice Examinations," *Psychological Reports*, Vol. 31, pp. 667-673.

Jacobs, Stanley S. (1972) "Answer Changing on Objective Tests: Some Implications for Test Validity," *Educational and Psychological Measurement*, Vol. 32, pp. 1039-1044.

Jarrett, R. F. (1948) "The Extra-Chance Nature of Changes in Students' Responses to Objective Test-Items," *The Journal of General Psychology*, Vol. 38, pp. 243-250.

Matthews, C. (1929) "Erroneous First Impressions on Objective Tests," *Journal of Educational Psychology*, Vol. 20, pp. 280-286.

Mueller, Daniel J. and Virginia Wasser, (1977) "Implications of Changing Answers on Objective Test Items," *Journal of Educational Measurement*, Vol. 14, No. 1, pp. 9-13.

Reile, Patricia J. and Leslie J. Briggs (1952) "Should Students Change Their Initial Answers on Objective-Type Tests?: More Evidence Regarding An Old Problem," *Journal of Educational Psychology*, Vol. 43, pp. 110-115.

Reiling, Eldon and Ryland Taylor (1972) "A New Approach to the Problem of Changing Responses on Multiple Choice Questions," *Journal of Educational Measurement*, Vol. 9, No. 1, pp. 67-70.

Samson, G. (1985) "Effects of Training in Test-Taking Skills on Achievement Test Performance: A Quantitative Synthesis," *The Journal of Educational Research*, Vol. 78, No. 5, pp. 261-266.

Vidler, Derek and Richard Hansen (1980) "Answer Changing on Multiple-Choice Tests," *Journal of Experimental Education*, Vol. 49, No. 1, pp. 18-20.