A Study of Answer Changing Behavior in MC Tests: The Effect of the Academic Course, Field of Study, Gender and Teachers' Attitude

Alireza Ahmadi

Shiraz University, Iran

arahmadi@shirazu.ac.ir

Abstract

This study investigated the answer changing behavior of the Iranian university students and its relationship to the academic course, gender, and field of study. 362 students at the University of Isfahan took part in the study. 76 English language teachers were also surveyed for their attitude toward answer changing on MC tests. The results indicated that 42.11% of the teachers had a negative attitude toward answer changing as they believed that the first answer selected is intuitively the best one and should not be changed. The results also indicated that answer changing would basically have positive effects on the total test scores. It was found that 63.09% of the answer changes were beneficial (from Wrong to Right choices). In contrast, 15.18% of the answer changes were negative (from Right to Wrong options), and 21.73% had no effect on the test scores (from Wrong to Wrong options). As for the effect of gender, no significant difference was found between males and females. Concerning the effect of academic course, the only significant difference lay in the number of Right to Wrong changes. Finally, it was found that field of study would make a difference only when Right to Wrong changes were considered.

Keywords: Answer Changing, Academic Course, Field of Study, Gender, Iranian English Teachers

1. Introduction

One of the most-widely-used tests to evaluate one's achievement has been the multiple choice (MC) test. Quite a few strategies have been recommended for successful performance on this test format; however, some of these strategies seem to be misbeliefs which are based on common sense rather than empirical investigations (Stough, 1993). One of these common misbeliefs which has attracted researchers for empirical investigation is the idea that students should not change their answers on MC tests because the first answers selected are intuitively the best answers and changing them leads to an increased number of wrong answers rather than improvement on one's score (Bauer, Kopp and Fischer, 2007). Sometimes even faculty members advise their students not to change answers (Kusler, 1988). However, studies conducted on the validity of this idea have basically rejected its value (e.g., Benjamin, Cavell, & Shallenberger, 1984). These studies indicate that most of the answer changers are point gainers (Al Hamly & Coombe, 2005). However, the idea that answer changing leads to losing points still prevails in some EFL contexts.

2. Literature Review

Answer changing has been the subject of much research. Most of the findings in this regard indicate that it can be considered as a positive test taking strategy in that most of the changes made are from incorrect to correct answers (Al-Hamly & Coombe, 2005; Balance, 1977; Bauer, Kopp & Fischer, 2007; Foote & Belinky, 1972; Friedman-Erickson, 1994; Geiger, 1991a; Heidenberg & Layne, 2000; Lynch & Smith, 1972; Waddell & Blankenship, 1994). Al Hamly and Coombe (2005) state that the most consistent findings of the studies conducted on answer changing can lead one to conclude that: only a small percentage of

items undergo changing; most of the test takers tend to change some of their answers; the change is mostly from the wrong answer to the right one; and consequently most of the answer changers are point gainers.

Answer changing has been studied for its relationship with various factors. A number of these studies focusing on gender have indicated that overall females make more answer changes than males (Bath, 1967; Foote & Belinky, 1972; Reile & Briggs, 1952; Skinner, 1983); however, males make more wrong to right answer changes than females (Pascale, 1974; Reile & Briggs, 1952; Skinner, 1983). The reverse is also reported by certain studies like Al-Hamly and Coombe (2005) who studied MCQ answer changing for Arab students. The results indicated that females made more wrong to right changes (62%) than males (52%).

Other studies report no relationship between gender and answer changing. For instance Heidenberg and Layne (2000) in a large scale study investigated the answer changing patterns in two graduate-level courses. Eight years of data including 1,819 subjects and 123,548 items indicated that approximately two-thirds of these changes were positive. The results revealed no significant difference between female and male students. They concluded that the rationale behind answer changing was related to the ability level of the student; that is, answer-changing frequency increased as the ability level decreased. The number of positive changes also decreased significantly by a decrease in the ability level. Formerly, similar results were found by Mercer (1979), Payne (1984), Penfield and Mercer (1980), and Ramsey et al. (1987).

Gender is not the only factor which has been studied in its relation to answer changing. Other studies have focused on factors like cognitive styles (Friedman & Cook, 1995); item difficulty (Ramsey et al., 1987; Vidler & Hansen, 1980); ethnicity (Payne, 1984); and test anxiety (Green, 1981).

In spite of the fact that previous studies have basically lent support to the efficiency of answer changing, the idea still prevails in some contexts, particularly in the Iranian EFL context, that answer changing is not an effective strategy. Some high school teachers and sometimes even faculty members discourage their students from changing their initial answers on MC tests. As such, some students may have become reluctant of changing their answers even if when upon reflection their first answer seems to be wrong. This was the motive for the present study to check the validity of this idea in an Iranian EFL context. The role of gender, academic course, and field of study was also taken into account to this end.

The reason why these three factors were selected for the present study was as follows. Gender was selected as a possible determining factor in response changing because although there have been a good number of studies about gender, the results of this studies are not at all conclusive. The idea that response changing may differ from one course to another due to the nature of the materials taught in each course caused the researcher to take the role of the academic course into account (BA students in Iranian universities are required to get about 136 course credits to graduate. It should be mentioned that some courses lend themselves more to recitation than comprehension or analysis and thus the type of questions used to test them would differ from other courses. For example questions that appear in a test of grammar are related to specific structural points and sometimes to language exceptions which are required to be memorized by the learners whereas the questions that appear in some other courses like research methods may need understanding and analysis on the part of learners. This means that the answer changing behavior may vary from course to course because the questions are of different types. Field of study was also considered in this study since BA students majoring in English (Literature,

Translation, and TEFL) in Iran seem to be different from the students of other fields of study as they are required to pass a specific course about language testing. Students of other fields are not basically familiar with the principles of testing and measurement in general let alone testing language skills. As such, English majors may benefit more from the response changing behavior due to their familiarity with the testing principles and test taking strategies.

Furthermore, the fact that there has been no study at all on answer changing in Iran, to the best of researcher's knowledge, and only a few studies on the role of the academic course and field of study in other contexts calls for more studies in this regard.

3. Research Questions

This study aims at investigating the answer changing behavior and its possible relationship with gender, academic courses (subject matters), and field of study. It also tries to shed light on the Iranian English teachers' attitude toward answer changing. Hence the following questions are put forward:

1. What do Iranian EFL teachers think of answer changing on MC tests?
2. To what extent do Iranian students get engaged in changing their first answers on MC tests?
3. Is answer changing a beneficial test taking strategy for students?
4. Is there a relationship between answer changing and gender?
5. Is there a relationship between answer changing and the academic course?
6. Is there a relationship between answer changing and field of study? In other words, is there a difference between Iranian English (EFL) majors and non-English majors in terms of answer changing?

4. Method

4.1. Participants

A total of 362 Isfahan University students and 76 English language teachers took part in this study. The majority of the students were female (279 females and 83 males). This is the usual norm in the Iranian university classes. Today, more than 70% of the students in many fields are females. Students were selected from English and non-English majors, 245 and 117 respectively. Convenient sampling was employed in selecting English majors since all the classes available were chosen for the study. Non-English majors were taking the General English Course. Students from different majors could attend the general English classes and as such the classes were basically mixed embracing students from different fields of study. Out of more than 10 classes held for the General English Course, 3 classes were randomly selected for this study. Non-English majors were selected since the researcher wanted to investigate the possible relationship between field of study and answer changing; that is, whether English majors and non-English majors would indicate the same behavior toward answer changing? A summary of the participants' major, level, and course of instruction is provided in Table 1.

Table 1. Descriptive Statistics of the Participants: Course, Level, and Major

Participants	Non-English Majors		English Majors		
	Mostly Freshmen	Freshmen	Sophomores	Juniors	Seniors
Course	General English	Grammar	Reading	Research	Teaching Methods
Number	117	76	79	46	44

Seventy six English language teachers randomly selected from different schools in the provinces of Isfahan and Fars also took part in this study.

4.2. Instruments

4.2.1. General English Test

A General English Test was given to non-English majors at the end of the academic semester as their Final Exam. This test consisted of 60 MC items in the areas of vocabulary, grammar, and reading comprehension. The reliability and validity of the test had already been established by the University Language center. Usually, General English Tests are written and validated by a group of EFL experts working for that center.

4.2.2. English Major Final Exams

Four English major final exams were also used as data collection instruments. These MC tests, namely grammar, reading comprehension, research methods and teaching methods, were given to students at the end of the academic semester as achievement tests. The tests consisted of 80, 60, 40, and 70 items respectively. To make sure that the tests had validity, all the items in each test were meticulously analyzed for content validity. The reliability of the four tests was also estimated through Cronbach's Alpha. The reliability indices turned out to be satisfactorily high for all the tests; that is, .79, .78, .74, and .83 for the tests of grammar, reading comprehension, research methods and teaching methods respectively.

4.2.3. Phone Interviews

English teachers were asked through phone calls to express their ideas concerning the answer changing behavior and why they thought it would be worth (not) changing the answers. They were asked if they recommend their

students to change their first answers on a MC test whenever in doubt about those answers.

4.3. Data Collection Procedure

The data of this study came from the final exams given to English and non-English majors; namely, four MC tests of grammar, reading comprehension, research methods and teaching methods and one MC test of General English. The data about the English teachers' attitudes toward answer changing were also collected through individual phone calls. All the phone calls were made by the researcher himself to add to the validity of the data collected.

4.4. Scoring Procedure

Students were expected to mark their answers on special answer sheets. They were told to use only pens to indicate their correct answers with a cross mark. In case they decided to change an answer, they were instructed to indicate the change clearly by filling in (shading) the circle related to the wrong answer completely. All the answer sheets were hand-marked and the scores were recorded manually. Each answer sheet was carefully analyzed by the researcher to find the instances of answer changing. The direction of each change was also carefully studied and manually recorded. Three directions for answer changes were possible: 1. wrong to right, 2. right to wrong, and 3. wrong to wrong.

4.5. Data analysis

The data collected underwent several stages of analysis. At first the data were analyzed for the total number of participants who were engaged in answer

changing. The number and quality of changes made in relation to gender, academic courses, and field of study were also analyzed separately. Phone interviews with the teachers were also carefully transcribed and analyzed.

5. Results

This study tried to find answer to six questions concerning answer-changing behavior in MC tests. In what follows the results are presented for each research question separately.

Research question 1: What do Iranian EFL teachers think of answer changing on MC tests?

The results of the phone interviews indicated that 42.11% (32 out of 76) of the English language teachers believed that answer changing is not a logical strategy. In fact they believed that the first answer selected is intuitively the best one and students should not change their first answer even if they are in doubt about that answer.

Research question 2: To what extent do Iranian students get engaged in changing their first answers on MC tests?

The results also indicated that overall 71.55% of the participants were engaged in changing answers; that is, 259 out of 362 participants changed at least one of their answers. This is in line with the literature that reports a similar percentage for the total number of answer changes, e.g., Al-Hamly and Coombe (2005) found that 67% of the participants in their study were engaged in answer changing.

Research question 3: Is answer changing a beneficial test taking strategy for students?

It was also found that 63.09% of the answer changes were beneficial; that is, the changes were from wrong to right choices. In contrast, 15.18% of the answer changes were negative (from right to wrong options), and 21.73% had no effect on the test scores since the changes were from wrong to wrong options.

The findings overall were in line with the literature in that most of the changes made were from wrong to right answers. Al-Hamly and Coombe (2005) state that usually a percentage between 51 to 59% is reported in the literature for the changes made from wrong to right answers, e.g., Al-Hamly and Coombe (2005), report 57%; Bauer, Kopp, and Fischer (2007), 55%; Foote and Belinky (1972), 55%; Friedman-Erickson (1994), 56%; Geiger (1991a), 59.9%; Waddell and Blankenship (1994), 57%. Sometimes higher percentages are reported in this regard like what was found in this study (63.09%) or in previous studies such as, Balance (1977), 86%; and Lynch and Smith (1972), 68%. Also in a large scale study Heidenberg and Layne (2000) found that approximately two-thirds of the changes were from wrong to right answers.

Concerning the right to wrong changes the literature reports a range between 19 to 24% (Al-Hamly & Coombe, 2005). One reason for the low percentage of right to wrong changes in the present study (15.18%) could be the fact that except for one group, all the other groups were English majors. They may be said to have the specialized knowledge required for the successful performance on language tests; that is, BA students majoring in English in Iran are different from the students of other fields because they pass a specific course on language testing. This course is considered one of the major courses in BA especially for those students who want to sit for the MA Entrance Examination in TEFL. Students of other fields are not basically familiar with the principles of testing and measurement in general let alone testing language

skills. This is because they do not pass any specific course on testing or measurement. As such, it seems that English majors benefit more from the response changing behavior due to their familiarity with the testing principles and test taking strategies. Consequently this has led to a higher degree of wrong to right changes and to a lower degree of right to wrong changes than what is commonly reported in the literature. This seems logical since studies that have investigated answer changing on language tests with non-language majors have reported a lower degree of wrong to right and a higher degree of right to wrong changes than what was reported in this study. For instance Al-Hamly and Coombe who employed a language test (Michigan English Language Institute College English Test) with non-language majors report 44% for wrong to right, 19% for right to wrong, and 37% for wrong to wrong changes.

This justification could also be supported by comparing the changes made by language and non-language majors in the present study. Table 3 indicates that students who were from other fields of study like physics, chemistry, accounting, etc. had the lowest degree of wrong to right changes (56.35%) which is in line with the range reported in the literature but it's lower than that of the other groups whose major was English. English majors indicated a higher degree of wrong to right changes (higher than 62%). The same was true for right to wrong changes (26.19% for non-English majors versus 4% to 18% for English majors).

The study also indicated that 21.73% of the changes were from wrong to wrong options which means that the changes had no effect on the total test score. This is in line with the literature since a similar percentage is usually reported for such changes e.g., Bauer, Kopp, and Fischer (2007), report 20% in this regard; also Foote and Belinky (1972), report 24%; Geiger (1991a), 19%;

11

and Skinner (1983), 22.3%. Sometimes higher degrees are reported for this type of change e.g., Al-Hamly and Coombe who employed English Language Tests with non-language majors report 37% in this regard.

Research question 4: Is there a relationship between answer changing and gender?

As for the relationship between answer changing and gender, the results indicated that females had a higher tendency than males to change answers (72% versus 69.35% respectively). This finding supports the relevant literature (Bath, 1967; Foote & Belinky, 1972; Reile & Briggs, 1952; Skinner, 1983).

Table 2. Answer Changes by Gender

	Right to wrong		Wrong to Right		Wrong to Wrong		Total	
	P	F	P	F	P	F	P	F
Male	62%	102	16%	27	22%	36	100%	165
Female	59%	358	18%	112	23%	137	100%	607

However, concerning the overall advantage that test takers took from changing their answers, as depicted in Table 2, males were more benefited. This is in line with the previous studies (e.g., Pascale, 1974; Reile & Briggs, 1952; Skinner, 1983).

A point is worth mentioning here that the contradictory results sometimes reported in the literature could partly be due to the fact that some studies have only relied on simple descriptive statistics like frequency and percentage to study the answer changing behavior and no statistical tests have been employed to support the findings. Hence the absence of statistical tests and the fact that raw data in terms of frequency and percentage could be misleading, may account for some of the contradictory findings in the literature. The present study, however, has utilized statistical tests to compare the results. Therefore a chi-square test was run to see whether the observed differences between males

and females were significant or not. No significant difference was found in terms of the total number or direction of answer changes (x2= 0.30, p= 0.86). This is in line with Mercer (1979), Payne (1984), Penfield and Mercer (1980), and Ramsey et al. (1987).

Research question 5: Is there a relationship between answer changing and the academic course?

One of the factors which is generally overlooked in the studies of answer changing is to see whether this behavior happens similarly across different academic courses. The present study tried to shed light on this question as well. The number and direction of answer changes across the academic courses are presented in Table 3 below.

Table 3. Answer Changes by Academic Course

Course	Wrong to Right		Right to Wrong		Wrong to Wrong		Total	
	P	F	P	F	P	F	P	F
General Eng.	56.35%	142	26.19%	66	17.46%	44	100%	252
Grammar	63.64%	56	15.90%	14	20.46%	18	100%	88
Reading	66.67%	104	11.54%	18	21.79%	34	100%	156
Research	66.67%	64	4.17%	4	29.16%	28	100%	96
Teaching Methods	62.22%	112	17.80%	32	20.00%	36	100%	180

The Table indicates that most of the changes are from wrong to right options regardless of the academic course. In a large scale study conducted by Heidenberg and Layne (2000) it was indicated that overall about two thirds of the changes made in the previous studies of answer changing were from wrong to right answers. The conclusion they have reached is also true across different academic courses.

It is also shown in Table 3 that the lowest percentage for wrong to right changes is related to the General English Course and the highest to the reading comprehension and research courses. A chi-square test was run to see the

significance of these differences. It was found that answer changing differed significantly across the academic courses (x2= 31.07, p= 0.000).

One of the problems in interpreting chi-square tests is the determination of which cell or cells produce the statistically significant difference. Examination of percentages in the contingency Table and expected frequency Table can be misleading. In order to see where the difference is exactly located, one may compare the size of the standardized residuals to the critical values that correspond to an alpha of 0.05 (+/-1.96) or an alpha of 0.01 (+/-2.58). This is equivalent to testing the null hypothesis that the actual frequency equals the expected frequency for a specific cell. There can be 0, 1, 2, or more cells with statistically significant standardized residuals to be interpreted. Standardized residuals that have a positive value mean that the cell was over-represented in the actual sample, compared to the expected frequency, i.e., there were more subjects in this category than were expected. Standardized residuals that have a negative value mean that the cell was under-represented in the actual sample, compared to the expected frequency, i.e., there were fewer subjects in this category than were expected.

Using standardized residuals (Appendix A), the study found that only the General English and Research courses were the significant contributors to the chi-square relationship between the academic course and answer changing. Of course the significant difference only referred to the number of right to wrong changes for the General English Course which was much higher than what was expected, and also for the research methods which was lower than what was expected. No significant difference was found across different academic courses when the number of wrong to right or wrong to wrong changes were considered.

By itself, chi square helps only to decide whether the variables are independent or related. It does not tell us how strongly they are related. Part of the reason is that the sample size and table size have such an influence upon chi square. When chi square is adjusted it becomes the basis for assessing strength of relationship. Cramer's V, an adjusted test to the chi-square test, is a measure of the strength of relationship for tables larger than 2 x 2 which is quite popular in this regard. To this end, it was utilized in the present study and indicated a moderately strong relationship between the answer changing behavior and the academic course (v= 0.28).

Research question 6: Is there a relationship between answer changing and field of study? In other words, is there a difference between Iranian English (EFL) majors and non-English majors in terms of answer changing?

To find the relationship between the field of study and answer changing, all the English majors were considered as a single group and were then compared with non-English majors. The results of a chi-square test revealed a significant difference ($x2$= 20.61, p= 0.000). However, to see where the difference exactly lay, standardized residuals were compared. Here again it was found that field of study did not make a difference as far as wrong to right or wrong to wrong changes were concerned, whereas it made a significant difference when right to wrong changes were considered. It was indicated that the number of right to wrong changes for non-English majors was much higher than what was expected and for English majors was much lower than what was expected. Cramer's V indicated a moderate relationship between field of study and answer changing (v= .24).

6. Discussion

This study shed light on a number of questions related to answer-changing behavior. The first question was whether Iranian EFL teachers had a positive or negative attitude toward answer changing on MC tests. The study indicated that 42.11% had a negative idea toward answer changing. This demands due attention since many students may become deprived of the beneficial effect of answer changing due to the wrong idea that their teachers hold toward answer changing. They may discourage their students from making any changes to their first answers even if when upon reflection their first answers seem to be wrong.

The second question was concerned with the extent to which Iranian students got engaged in answer changing behavior. The present study found that overall 71.55% of the subjects were engaged in answer changing. This finding, though in line with the literature, is very surprising for the Iranian context. Most of the students were engaged in answer changing in spite of the fact that many might have had a negative idea toward answer changing since 42.11% of their teachers had a negative attitude toward answer changing and discouraged them from changing their answers. This needs further qualitative investigations. However, one probable reason for such a behavior may be the fact that students are not penalized for their wrong answers on MC tests they are given as their midterm or final exams. Thus they may be encouraged to guess the right answer and to change their answers more freely whenever they are not sure about their answers. This is against what is ordinarily done in the Iranian University Entrance Examination in which students are penalized for their wrong answers. They are discouraged from guessing the write answer or changing their first answer because their mind is obsessed with the negative mark they receive for their wrong answers.

The third question dealt with the direction of answer changing and whether it leads to beneficial effects. The study found that answer changing is a beneficiary behavior in contrast to what is commonly believed by Iranian teachers and students. The study indicated that overall 63.09% of the changes were from wrong to right answers, that is, most of the changes were beneficial. Only 15.18% of the changes were form right to wrong supporting the commonly-held misbelief that answer changing leads to negative effects because the first answer selected is intuitively the best. Also 21.73% of the answer changes were from wrong to wrong options having no effect on the total test score. These findings were in line with the literature (e.g., Al-Hamly & Coombe, 2005; Balance, 1977; Bauer, Kopp, & Fischer, 2007; Waddell & Blankenship, 1994; Heidenberg & Layne, 2000; Foote & Belinky, 1972; Friedman-Erickson, 1994; Geiger, 1991a; Lynch & Smith, 1972), though usually previous studies had reported a lower percentage for the wrong to right changes and a higher percentage for the right to wrong changes than what was found in this study. This was explained through the effect of field of study. The present study indicated that when the language test is taken by the language majors, higher percentage for wrong to right and lower percentage for right to wrong changes are found than when the language test is taken by non-language majors.

As for the forth question concerning the relationship between gender and answer changing behavior, the present study revealed no significant relationships. Although a slightly higher percentage of males made wrong to right changes than females (61.81% versus 58.91% respectively), the difference was not significant according to a chi-square test ($x2= 0.30$, $p= 0.86$). The same thing was true for right to wrong and wrong to wrong changes.

The fifth question looked for the probable relationship between the academic course and the number and direction of answer changes. The assumption was that the response changing behavior could vary based on the academic course because different courses lend themselves to different types of questions. Thus response changing would probably be more useful when dealing with MC questions that require comprehension and analysis. In contrast, it is not usually effective when questions which are somehow related to rote learning and memorization are at focus. Answer changing related to these questions is only effective if the memory serves the person right or if the person is lucky enough to have a positive change based on pure chance. This assumption is supported in this study only if the raw frequencies or percentages are considered (Table 3). The study indicated that the research and reading comprehension courses had the highest percentage of wrong to right changes. The research course also had the lowest percentage of right to wrong changes. Hence it can be said that the research course had the best results for response changing, and reading comprehension occupied the second rank. Grammar and teaching methods came next and the general English course lagged behind. This can be justified in line with the above-mentioned assumption because most of the questions in the test of research were those demanding comprehension and analysis. Also many reading comprehension questions were of the type that needed deep comprehension and analysis of the elements and their relationship in the text. Grammar and Teaching Methods included many examples of the questions requiring memorization and recitation.

Of course the fact that the General English course occupied the last position in response changing behavior is not in line with the assumption mentioned above, since the questions in this part were more or less similar to the questions that appeared in the Grammar test. This means that another

factor is playing a role here and the academic course cannot be a decisive factor. When the results are studied carefully, it can be said that this difference may be due to the effect of field of study, since all the students taking the General English course were non-English majors whereas those taking the other four courses were English majors. Another possible explanation could be the effect of level of proficiency because ordinarily those non-English majors taking the General English Course hold a lower mastery over the language skills than their English major counterparts.

The last question asked for the relationship between field of study and answer changing; that is, whether English majors would be different from non-English majors in terms of their answer changing behavior? The study concluded that field of study could make a significant difference only in case of right to wrong changes. The number of right to wrong changes for non-English majors was higher than expected and for English majors was lower than expected. That is probably because English majors have more specialized knowledge about language test taking strategies (taking a specific course on language testing) or because they are more motivated than non-English majors in learning English. For non-English majors, language is considered as a peripheral course studied only in one term and hence it may not receive due attention on their part since other courses related to their field of study are considered more important. Furthermore, since they do not pass a course on testing they may not be familiar with testing principles and language test taking strategies. Overall, this may mean that English language majors make wiser changes on MC language tests than non-language majors.

7. Conclusions

This study indicated that the traditional common (mis)belief that is still held by many language teachers and practiced by learners, specifically in an Iranian EFL context, is not well-grounded. The study found that answer changing overall brings about gains in one's total test score. The study also concluded that gender, academic course, and field of study do not play any significant roles in this regard since the participants in this study regardless of their gender, field of study or the course they were taking benefited from answer changing. As such, EFL students should be encouraged to get themselves involved in changing answers providing that through reflection their first answer seems to be wrong. It could be mentioned in line with Bauer, Kopp, and Fischer (2007) that one suggested consequence of such answer changing studies is that examinees should be informed about this misbelief in the hope that this prejudice be eradicated.

8. Limitations of the Study

The following limitations should be considered when one is trying to generalize the findings of the present study.

The fact that different groups took different tests in the present study, or that they were from different years of study (roughly meaning different proficiency levels) might have blurred the results of the study since the effect of academic courses might have been confounded with the effect of level of proficiency. However, it should be underlined that some previous studies have basically shown no relationship between the level of proficiency and answer changing behavior (e.g., Mercer, 1979; Pascale, 1974; Ramsey et al., 1987). Furthermore, the results of this study indicated a high level of similarity among

the English majors from different years of study and the only significant difference was found between the English majors and non-English majors, leading one to conclude that field of study had played a more significant role in this study than the level of proficiency.

Answer changing may vary across test items depending on factors such as the difficulty level of the task and nature of the given MC options. These factors were not considered in the present study and hence due caution is needed in generalizing the findings.

9. Pedagogical Implications

The results of this study could have implications for language teachers and students. The fact that most of the changes made by students were from wrong to right answers calls for a more realistic view toward answer changing on MC tests. This is especially true in an Iranian EFL context, as this study indicated that 42.11% of the Iranian English teachers still stick to the traditional belief that the first answer selected in a MC test is intuitively the best one and therefore answer changing leads to reduction in one's overall test score. This study recommends that teachers consider answer changing as a useful technique for both males and females and encourage their students to change their answers whenever upon reflection their first answer seems to be wrong. Bauer, Kopp, and Fischer (2007), state that students encouraged and instructed to change answers when in doubt, will heed the advice and change more answers than students not so instructed.

The results of the study make one think of the nature of different courses and the role it may have in answer changing. A trend was found in this study for the effect of field of study and the academic course on answer changing, though

the results were not significant in this regard. It seemed that answer changing would be more beneficial when MC questions needed a kind of analysis. Also, the theoretical knowledge of testing principles and strategies seemed to play a role in this regard, as those students who lacked this knowledge had fewer beneficial changes on MC items.

The results could also be of use to those involved in practicing or developing MC tests. This study indicated that answer changing can be considered as an effective test-taking strategy. Thus it's reasonable that those teaching or practicing test-taking strategies consider answer changing as an effective test-taking strategy and provide students with a sound basis for answer changing which is grounded in research.

## 10. Recommendations for Further Research

Although there is a good deal of research on answer changing, very few studies may be found which have specifically focused on this issue in EFL contexts. As such, great need is felt in EFL contexts for similar studies to be carried out before one can make sound conclusions. A future line of research may hence replicate the same study in EFL contexts or focus on answer changing and its relation to cognitive styles, affective factors, etc. For example one may focus on how answer changing varies across individuals. This study indicated that answer changing overall is an effective strategy. But it may vary from person to person due to the individual characteristics. Thus determining which characteristics are more influential in this regard would be a fertile area for research.

A different line of research could focus on the efficacy of answer changing as a test taking strategy and compare it to other test taking strategies under different conditions.

The studies conducted so far have focused on the number and direction of changes made. Studies in this regard have indicated that about two thirds of the changes are from wrong to right. However, more studies could probably be carried out focusing on the causes for different changes. They may compare positive changes with negative changes, and hopefully come across guidelines that can increase the efficiency of this technique.

Finally as the results of this study indicated a trend that favored answer changing when the MC question demanded a kind of analysis or deeper comprehension in answering the item than when it required remembering a memorized concept. This idea can specifically be put into investigation by including non-language courses that are analytical in nature, like mathematics, and comparing them with language courses.
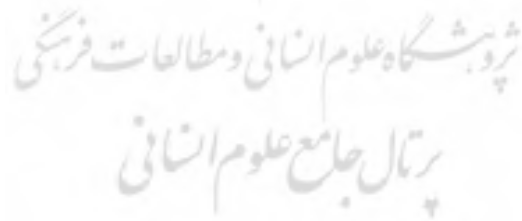
References

Al-Hamly, M., & Coombe, C. (2005). To change or not to change: Investigating the value of MCQ answer changing for Gulf Arab students. *Language Testing, 22* (4), 509-531.

Balance, C. (1977). Students' expectations and other answer changing behavior. *Psychological Reports, 41*, 163-166.

Bath, J. (1967). Answer-changing behavior on objective examinations. *The Journal of Educational Research, 6*, 105-107.

Bauer, D., Kopp, V., & Fischer, M. (2007). Answer changing in multiple choice assessment change that answer when in doubt-and spread the word. *BMC Medical Education*, 7:28. doi:10.1186/1472-6920-7-28.

Benjamin, L. T., Cavell T. A., & Shallenberger, W. R. (1984). Staying with initial answers on objective tests – Is it a myth?. *Teaching of Psychology, 11*, 133-141.

Foote, R., & Belinky, C. (1972). It pays to switch? Consequences of changing answers on multiple-choice examinations. *Psychological Reports, 31*, 667-673.

Friedman, S., & Cook, G. (1995). Is an examinee's cognitive style related to the impact of answer changing on multiple choice tests?. *Journal of Experimental Education, 63*, 199-213.

Friedman-Erickson, S. (1994). *To change or not to change: The multiple choice dilemma*. Paper presented at the annual institute of the American Psychological Society on the Teaching of Psychology, Washington, DC.

Geiger, M. (1991a). Changing multiple choice answers: A validation and extension. *College Student Journal, 25*, 181-186.

Green, K. (1981). Item-answer changes on multiple choice tests as a function of test anxiety. *Journal of Experimental Education, 49*, 225-228.

Heidenberg, A. J., & Layne, B. H. (2000). Answer changing: A conditional argument. *College Student Journal, 34*(3), 440-450.

Kusler, D. (1988). Teaching students how to take an examination. *Journal of Emergency Nursing, 14*, 190-191.

Lynch, D. O., & Smith, B. C. (1972, April). *To change or not to change item answers when taking tests: Empirical evidence for test takers.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Mercer, M. (1979, April). *Answer changing and students' test scores.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Pascale, P. (1974). Changing initial answers on multiple choice achievement tests. *Measurement and Evaluation in Guidance, 6*, 236-238.

Payne, B. D. (1984). The relationship of test anxiety and answer-changing behavior: An analysis by race and sex. *Measurement and Evaluation in Guidance, 16*(4), 205-210.

Penfield, D., & Mercer, M. (1980). Answer changing and statistics. *Educational Research Quarterly, 5*, 50-57.

Ramsey, P., Ramsey, P. P., & Barnes, M. J. (1987). Effects of student confidence and item difficulty on test score gains due to answer changing. *Teaching of Psychology, 144*, 206-210.

Reile, P., & Briggs, L. (1952). Should students change their initial answers on objective-type tests? More evidence regarding an old problem. *Journal of Educational Psychology, 43*, 110-115.

Skinner, N. (1983). Switching answers on multiple-choice questions: Shrewdness or shibboleth?. *Teaching of Psychology, 14*, 241-242.

Stough, L. (1993, April). *Research on multiple-choice questions: Implications for strategy instruction.* Paper presented at the annual convention of the Council for Exceptional Children, San Antonio, TX.

Vidler, D., & Hansen, R. (1980). Answer changing on multiple-choice tests. *Journal of Experimental Education, 49*, 18-20.

Waddell, D. L., & Blankenship, J. C. (1994). Answer changing: A meta-analysis of the prevalence and patterns. *The Journal of Continuing Education in Nursing, 25*(4), 155-158.