

Changing Answers in Multiple-Choice Exam Questions: Patterns of TOP-Tier Versus BOTTOM-Tier Students in Podiatric Medical School

Grant Yonemoto¹, Milad Kashani², Erica B Benoit³, Jeffrey J Weiss⁴ and Peter Barbosa⁵

¹Yale Podiatric Medicine and Surgical Residency Program, Yale New Haven Health, New Haven, CT, USA.

²Swedish Medical Center, Seattle, WA, USA. ³School of Health and Sciences, Universidad del Sagrado

Corazon, San Juan, PR. ⁴Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁵American

University of the Caribbean School of Medicine, Cupecoy, Sint Maarten.

Journal of Medical Education and
Curricular Development
Volume 10: 1–7
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23821205231179312



ABSTRACT

OBJECTIVES: Students often express uncertainty regarding changing their answers on multiple choice tests despite multiple studies quantitatively showing the benefits of changing answers.

METHODS: Data was collected from 86 first-year podiatric medical students over one semester for the course of Biochemistry, as shown in electronic testing data collected via ExamSoft's® Snapshot Viewer. Quantitative analysis was performed comparing frequency of changing answers and whether students changed their answers from incorrect-to-correct, correct-to-incorrect, or incorrect-to-incorrect. A correlation analysis was performed to assess the relationship between the frequency of each type of answer change and class rank. Independent-sample *t*-tests were used to assess differences in the pattern of changing answers amongst the top and bottom performing students in the class.

RESULTS: The correlation between total changes made from correct-to-incorrect per total answer changes and class rank yielded a positive correlation of $r=0.218$ ($P=.048$). There was also a positive correlation of $r=0.502$ ($P<.000$) observed in the number of incorrect-to-incorrect answer changes per total changes made compared to class rank. A negative correlation of $r=-0.382$ ($P<.000$) was observed when comparing class rank and the number of changed answers from incorrect-to-correct. While most of the class benefited from changing answers, a significant positive correlation of $r=0.467$ ($P<.000$) for percent ultimately incorrect (regardless of number of changes) and class rank was observed.

CONCLUSION: Analysis revealed that class rank correlated to likelihood of a positive gain from changing answers. Higher ranking students were more likely to gain points from changing their answer compared to lower ranking. Top students changed answers less frequently and changed answers to an ultimately correct answer more often, while bottom students changed answers from an incorrect answer to another incorrect answer more frequently than top students.

KEYWORDS: multiple-choice exam questions, medical education, assessment, TOP-Tier students, BOTTOM-Tier students

RECEIVED: June 22, 2022. **ACCEPTED:** May 15, 2023.

TYPE: Original Research Article

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was completed in part with the support of grants from the US Department of Education P120A190043 and P031C210100.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Peter Barbosa, American University of the Caribbean School of Medicine, #1 Jordan Road, Cupecoy, Dutch Lowlands, Sint Maarten. Email: peter.barbosa@aucmed.edu

Introduction

Despite multiple studies conducted over the years on the benefits of changing answers in multiple-choice tests, the old adage that a student should not change their initial answer still exists.^{1–6} Some students and faculty believe that changing answers is detrimental to the overall performance on a multiple-choice exam and that student anxiety during an exam affects judgment when reviewing previously answered questions.^{5,6} Even though studies conducted across many different health fields spanning from nursing, dental, and medical students have shown that overall, it is beneficial and recommended that students change their answer after careful reevaluation of the question.^{4,6–11}

In 1984, Benjamin et al analyzed 33 studies from 1928 to 1983 and concluded that a majority of answer changing is beneficial to the students despite the widespread belief to the

contrary.⁸ Since Benjamin's study, many studies have substantiated this conclusion, quantifying the benefit of changing ones answer across many different medical fields of study. A study done on US Army Anesthesia Nursing students showed that of the answers that were changed, 72% of the answers were changed from wrong to right (IC) while only 20% of the answers were changed from right to wrong (CI).¹² This benefit is also seen in a study that consisted of medical students in Germany indicating that there is a clear-cut improvement from incorrect to correct (IC) on their initial answer change. The same study also showed that there was a high correlation between the number of answers changed and number of points gained.⁹ The benefits of changing answers are shown to benefit all students in the class regardless of demographics or class rank.¹³ Another study performed on baccalaureate nursing students showed a significant positive score improvement



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

with changing answers in all tiers of students regardless of previous class performance.¹⁴

Although numerous studies show the benefit of changing answers, student behavior during examinations does not reflect this knowledge.¹⁵ German medical students educated about the benefits of changing answers showed a significant increase in the number of answers changed and overall better performance on the test when compared to the non-educated group.¹⁶ Students that may have been misled to believe that staying with the initial answer is more beneficial, often state that test anxiety and previous test experience influence their hesitation to change answers.^{5,6,17} Students often feel a lack of confidence and have the general feeling of always changing the right answer to a wrong answer (CI) based on previous experiences.^{9,18,19} All these studies also point to a critical element in assessment and higher education: not all students are the same, and not all solutions will be equally effective for diverse populations. Many of the studies referenced above tend to reach conclusions generalized for the entirety of their cohorts, not considering multifactorial elements presented by non-traditional students, various socioeconomic student populations, racial and ethnic diversity, and quality of K12 preparation, to mention a few. Is it possible that different subgroups of students would present different patterns of answering multiple-choice questions?

With the shift from paper-based multiple-choice exams to electronic computer-based exams, there has been no drop-off in benefits to switching answers. Electronic test taking data on first year dental medical students showed that 99.4% of the students saw a net benefit of at least 2 correct questions from changing their answers over the span of 7 exams in 2 different courses.¹¹ The use of computer-based examinations has allowed investigators to examine answer changes without previously used subjective measures and has allowed more in-depth analysis on individual test question performance.^{6,11,20}

The purpose of this study is to validate the previous findings on the benefits of changing answers in podiatric medical school examinations and to investigate if the benefits of changing answers are uniform across class rank with a specific focus on TOP-Tier versus BOTTOM-Tier podiatric cohorts.

Methods

This study was approved by New York College of Podiatric Medicine's (NYCPM) Institutional Review Board in July 2017. Informed consent was waived off by Institutional Review Board since students agree to the use of unidentifiable education data for research purposes at registration. Participants consisted of 86 first-year students in the Doctorate of Podiatric Medicine program and enrolled in the fall semester Biochemistry course in 2016. Inclusion criteria consisted of having a good-standing student status as a first-year student at NYCPM during the fall 2016; exclusion criteria included students who dropped out of the program prior to the

culmination of the Biochemistry course. This educational original research study took place in Manhattan, New York.

Exams were administered on school-issued iPads with ExamSoft's® software pre-installed. Any students who required accommodations or additional time were granted within the guidelines set forth by the American Disabilities Act.²¹ Using ExamSoft's® Snapshot Viewer feature, data was collected on all students enrolled over a span of 6 biochemistry examinations from August to December 2016. Incorporating electronic tracking and analysis of test taking behavior allowed for non-subjective bias seen in many previous non-computer studies. Of the 86 students who started the year, 84 of the students went on to complete the Biochemistry course. Each Biochemistry exam was an exam based on the progressive and cumulative knowledge learned thus far in the course and weighted equally towards final course grade calculations. The exams consisted of multiple-choice questions with one correct answer and 4 distractors, for a total of 5 choices. Students were allocated 1 min and 10 s per question and were permitted to return to any question during the exam period.

Statistical analysis

Quantitative analysis was done on the class, comparing frequency of changing answers and whether students changed their answers from incorrect to correct (IC), correct to incorrect (CI), incorrect to incorrect (II), or correct to correct (CC). The final evaluation was done by observing and comparing the initial answer choice to the final answer choice. Total number of changes were calculated by evaluating each time a student changed their answer between the student's initial and final answer choices. Analysis was then performed to compare each type of initial and final answer choice and the total number of changes observed.

Class ranking was determined by adding total number of ultimately correct answers and dividing by the total number of questions throughout the 6 exams. The highest performing student was ranked as 1 with the lowest ranking student ranked 84. Using IBM SPSS® version 24 software, a Spearman's rank order correlation was completed to evaluate the correlation between class rank in descending order and each category. Independent-sample *t*-test was used to compare groups. A 95% confidence level was used in this study.

Results

Exam results were collected from all 86 members of the Class of 2020 with data only being evaluated for members of the class that completed all 6 Biochemistry exams using the ExamSoft's® Snapshot Viewer feature. In total, 84 students completed all 330 questions across 6 exams which yielded a total of 27 390 questions analyzed. The mean cumulative exam score for the class was 79.4% (SD = 10.2%) with a mean of 22.43 (SD = 15.69) total changes per student. All students changed an answer at least once throughout all exams.

Quantitative analysis was performed on the aggregated data across 6 exams for each student. The net positive gain seen on test scores was calculated by aggregating the total amount of changes that resulted in a correct answer and comparing it to the total number of changes that resulted in an incorrect answer. The correlation between total changes made from correct to incorrect (CI) per total answer changes and class rank yielded a positive correlation of $r=0.218$ ($P=.048$) (Figure 1). There was also a positive correlation of $r=0.502$ ($P<.000$) observed in the number of incorrect to incorrect (II) answer changes per total changes made compared to

class rank (Figure 2). A negative correlation of $r=-0.382$ ($P<.000$) was observed when comparing class rank and the number of changed answers from incorrect to correct (IC) (Figure 3). Overall, 60% of students ultimately had a positive net benefit from changing answers. While a majority of the class benefited from changing answers, a significant positive correlation of $r=0.467$ ($P<.000$) for percent ultimately incorrect (regardless of number of changes) and class rank was also observed (Table 1).

To examine the extremes TOP-Tier versus BOTTOM-Tier, analysis was done on the top and bottom 10 students of

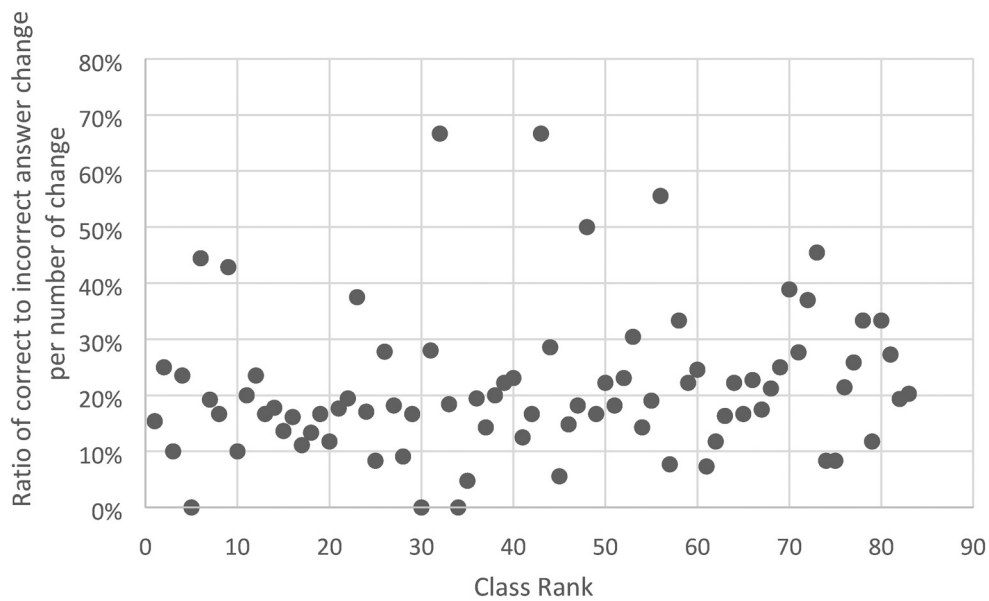


Figure 1. Correlation between the number of correct to incorrect (CI) answer changes per number of changes and class rank ($\rho=0.218$; $P=.048$).

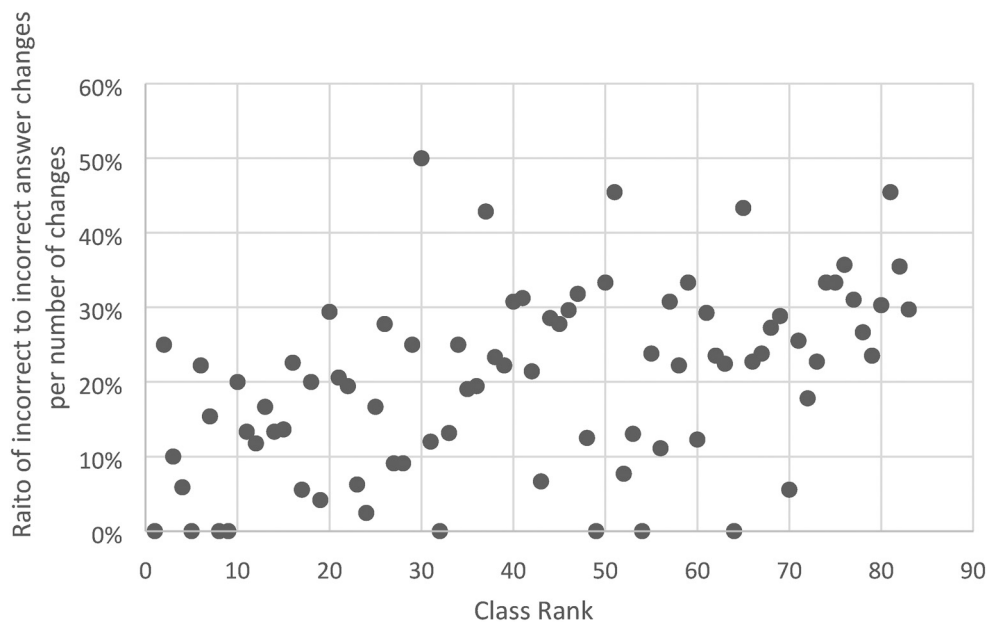


Figure 2. Correlation between the number of incorrect to incorrect (II) answer changes per number of changes and class rank ($\rho=0.502$; $P<.000$).

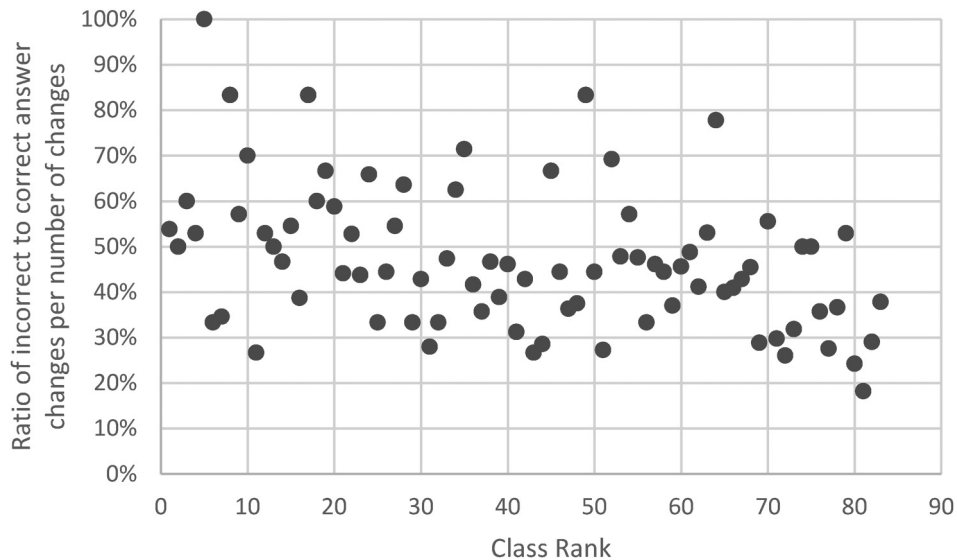


Figure 3. Correlation between the number of incorrect to correct answer changes per number of changes and class rank ($\rho = -.382$; $P < .000$).

Table 1. Spearman rank correlation analysis.

	Ratio of correct to incorrect (CI) per number of changes and class rank	Ratio of incorrect to incorrect (II) per number of changes and class rank	Ratio of incorrect to correct (IC) per number of changes and class rank	Ratio of total incorrect per number of changes and class rank
R	0.502	0.218	-0.382	0.467
P-value ($P < .05$)	<.000	.048	<.000	<.000

the class, comparing frequency of changing answers from incorrect to correct (IC), correct to incorrect (CI), or, incorrect to incorrect (II). The mean final test score for the top 10 students was 93.88% (SD = 2.42%) correct and 61.09% (SD = 5.83%) for the bottom 10 students. Top students changed answers less frequently with a mean of 10.9 answer changes compared to bottom students with 30.4 ($P = .017$) changes (Figure 4). Top students changed answers from an incorrect choice to a different incorrect choice (II) only 9.8% of the time which was lower when compared to changes observed in bottom students at 29.9% ($P < .000$). Top students also changed the answer from an initial incorrect choice to the correct answer (IC) more frequently (59.5%) compared to bottom students 38.7% ($P = .023$) (Table 2).

Discussion

While Saadat et al conclude that “The theory that students should trust their first instinct and stick to their initial answer on a multiple choice test is a myth that should be discarded,”⁵ our findings indicate that this conclusion may be proper for some students but not for others. Our study is consistent

with similar studies and suggests that overall, podiatric medical students benefit from changing test answers, but this holds true only for those who are top performers. Of the 84 students studied, every student changed an answer at least once over the 6 exams and 60% of students had an average net positive gain in exam scores as a result of changing answers. We observed that a mean of 6.25% of the questions were changed from their initial response, which is in line with previously published studies.^{8,11}

In contrast to what has been reported in other cohorts of students in the health sciences, the benefits of changing answers are not distributed equally across all podiatric students. A positive correlation of $r = 0.467$ ($P < .000$) between the number of ultimately incorrect answers after changing answers and class rank (Table 2) suggest that as overall performance decreases, there is a greater probability of selecting an ultimately incorrect answer after an answer change. These observations are also supported by a significant positive correlation comparing answer changes from an incorrect answer to another incorrect (II) answer per change and class rank (Table 2). When comparing the number of answer changes from correct to incorrect (CI) per change and class rank, lower performing students are more likely to change their answers from correct to incorrect (CI) thus hurting their overall grade.

There are several study limitations. Since the sample size of 84 was determined based on the number of students registered in the course, a sample size/power analysis was not performed, which represents a study limitation. Additional study limitations were in the areas of sample selection and duration of the longitudinal design. Only one group of students and one single subject were included in the study: biochemistry students in the Fall 2016. If multiple biochemistry cohorts were analyzed, or multiple courses were included,

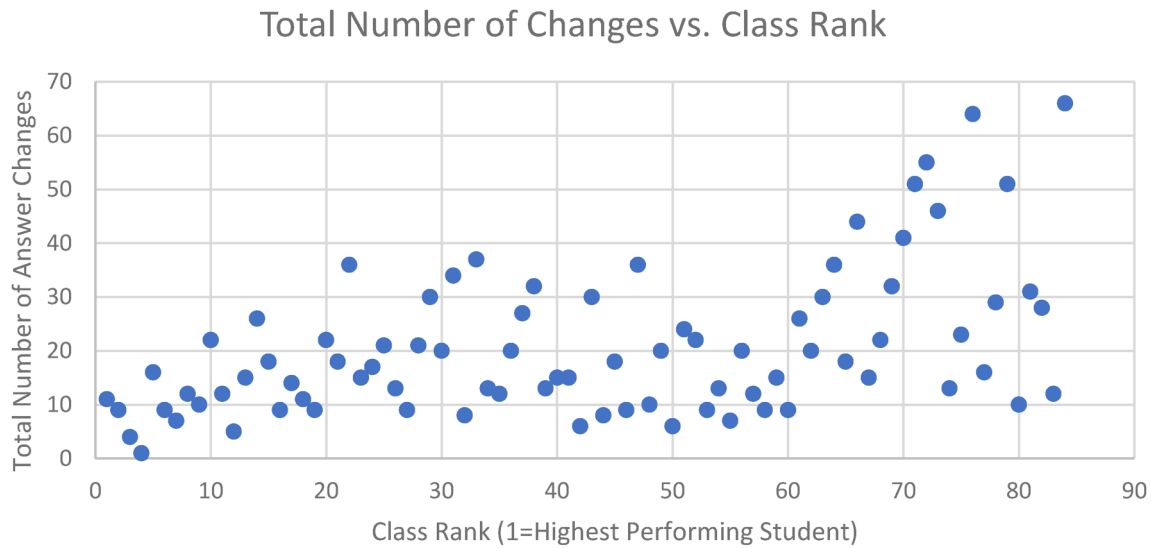


Figure 4. Total number of answer changes versus class rank.

Table 2. Numbers and analysis of top 10 and bottom 10 students in the course ranked by overall score.

	Mean total score (SD)	Mean total changes (SD)	Mean percent incorrect to correct (IC) (SD)	Mean percent incorrect to incorrect (II) (SD)	Mean percent ultimately correct (SD)	Mean percent ultimately incorrect (SD)
Bottom 10 students	61.09% (0.024)	30.40 (19.24)	38.72% (0.167)	29.96% (0.096)	41.46% (0.165)	51.72% (0.142)
Top 10 students	93.39% (0.058)	10.90 (6.97)	59.52% (0.205)	9.85% (0.101)	64.77% (0.187)	30.56% (0.192)
P-value	.000	.017	.023	.000	.008	.012

results would be more comprehensive. Additionally, the study was conducted only in the first semester of podiatric medical school. It is possible that the observed pattern could change if the same students were followed over the course of the 4-year curriculum.

Conversely, quantitative analysis showed a statistically significant negative correlation between the incorrect answers changed to correct answers (IC) per change compared to class rank. Intuitively, high performing students had a higher percentage of changing answers from incorrect to correct (IC) increasing their overall examination performance. Exam preparedness and understanding of the material increases the confidence a student will have in the chosen answer, making it less likely to change their initial answer as well as having a higher net gain from changing answers.^{10,11,22,23} Students who may have made a careless mistake or who remembered exam content as the exam progressed are also more likely to benefit from changing their answers.^{10,17,24}

When isolating TOP-Tier¹⁰ and BOTTOM-Tier¹⁰ students, Bottom students changed answer more frequently than Top students, which is consistent with previous studies.^{11,22-24} A significant difference was also observed in the frequency in

which Top students and Bottom students changed their answers from incorrect to correct (IC), as well as the frequency in answer changes from incorrect to incorrect (II) (Figure 3). Eight of the 10 Bottom-performing students did not have a net benefit from changing answers and lowered their overall performance. In contrast, there was only one Top-ten student who did not benefit from changing their answer.

One possible explanation for the observed trend is that low performing students who have less understanding of the material have less confidence in their answer choices. Cox-Davenport et al¹⁷ observed that students who had partial understanding of the material felt uncertain about their initial response and often narrowed down potential answers without feeling confident in their chosen answer choice. If a student does not fully understand the subject material and does not have confidence in his/her answers, they are more likely to change answers more frequently due to feeling anxious about his/her answer choices.^{5,6,10,17,22-24} Students who are not adequately prepared for the examination tend to be less confident in their answer choices and regret changing their answers believing that the answer was changed from a correct answer to an incorrect answer (CI).^{10,18,19} These observations are aligned with

previously described measures of self-monitoring, where high performing students showed greater positive differences.²⁵

Anxiety over final answer choices can also be caused by question stems that are worded in a way that is confusing or ambiguous, making it more likely for students to change their answers.¹⁰ Confusion with the clear comprehension of question stems, in addition to inadequate knowledge of the exam material, can lead to students having lower confidence in their initial answer and can cause a student to doubt their answer choice regardless of it being correct.^{22,23,26,27} Careful examination of the question and critical thinking can narrow down answer choices and reduce the anxiety a student may have while decreasing the number of careless mistakes.²⁸

Recognizing diversity and stratification in medical students and medical education is essential. Some go as far as to suggest that there are tiers of medical schools based on their ranking.^{29,30} In addition to rankings, elements of diversity, equity, and inclusion play a role in students' performance and test-taking behaviors.³¹ Non-traditional students bring an additional layer to this complexity of factors affecting test-taking performance.³² Our findings are relevant and significant to this medley of medical student types. Our results support the idea that different interventions, strategies, and approaches could be implemented for lower-performing students in medical schools. The use of computer-based exams allowed for unbiased analysis of a true answer choice change as well as allowing for a larger sample size. Every student analyzed took the exams on their school-issued iPads. All changes were assessed from data collected on ExamSoft's® software, and we were unable to distinguish the reason why students changed their answers. Prior knowledge of a study, like the one conducted, could have influenced student decision making and skewed results. Continued investigation is also needed to assess if changing one's answer is still beneficial in other course subjects, both in the pre-clinical and clinical courses.

Conclusions

Our study showed that overall, podiatric medical students benefited from changing their answers, which support findings published in previous studies among healthcare students. While changing one's answer may be beneficial when looking at the entire class, there was a notable difference in what type students had the greatest benefit from changing their answers.

Lower performing podiatric students changed their answers with increased frequency compared to top performing podiatric students. In contrast to other healthcare student cohorts, lower performing podiatric students also had a greater likelihood of their answer changes being ultimately marked incorrect. For most of these students, changing their answer was not beneficial to their overall grade.

Our primary finding supports the idea that caution should be taken when giving generalized advice to students about the

benefits of changing initial answers. Lower ranking students should carefully review questions in which they are unsure and use critical reasoning to narrow down potential answer choices before changing their answers. Future efforts and research should be examined and developed about how educational systems could improve this pattern observed among BOTTOM-Tier students.

Acknowledgements

This work originated at New York College of Podiatric Medicine; the authors express their deep gratitude for the support received. The authors of this paper would like to thank James Henkel, PhD (Associate Dean, School of Pharmacy and Physician Assistant Studies) at the University of Saint Joseph, for his instructional insights about the specific features of the ExamSoft software utilized in this analysis. The authors of this paper would also like to thank Gabriela Santiago Resto at Universidad del Sagrado Corazón for her contribution in formatting and submission of this article.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Geiger MA. Changing multiple-choice answers: do students accurately perceive their performance? *J Exp Educ*. 1991;59(3):250-257.
2. Gaskins S, Dunn L, Forte L, Wood F, Riley P. Student perceptions of changing answers on multiple choice examinations. *J Nurs Educ*. 1996;35(2):88-90.
3. Foote R, Belinky C. It pays to switch? Consequences of changing answers on multiple-choice examinations. *Psychol Rep*. 1972;31(2):667-673.
4. Merry J, Elenchin M, Surma R. Should students change their answers on multiple choice questions? *Adv Physiol Educ*. 2021;45(1):182-190.
5. Saadat S, Noori M, Ghazizadeh Hashemi AH, Mousavi bazaz N, Babakhanian M, Montazeri K, Hadem A. Students' challenge in answer-changing on multiple-choice exams; doubting the answer or not? A systematic review. *Med Edu Bull*. 2020;1(2):104-111.
6. Bultas MW, Schmuke AD, Rubbelke C, Jackson J, Moran V. Assessment of answer-changing behavior of nursing students and faculty recommendations. *J Nurs Educ*. 2021 Jun;60(6):324-328.
7. Geiger MA. On the benefit of changing multiple-choice answers: student perception and performance. *Education*. 1996;117(1):108.
8. Benjamin LT, Cavell TA, Shallenberger WR. Staying with initial answers on objective tests: is it a myth? *Teaching of Psychology*. 1984;11(3):133-141.
9. Fischer MR, Herrmann S, Kopp V. Answering multiple-choice questions in high-stakes medical examinations. *Med Educ*. 2005 Sep;39(9):890-894.
10. George TP, Muller MA, Bartz JD. A mixed-methods study of prelicensure nursing students changing answers on multiple choice examinations. *J Nurs Educ*. 2016 Apr;55(4):220.
11. Pagni SE, Bak AG, Eisen SE, Murphy JL, Finkelman MD, Kugel G. The benefit of a switch: answer-changing on multiple-choice exams by first-year dental students. *J Dent Educ*. 2017 Jan;81(1):110.
12. Johnson D, Anderson S, Johnson S. To change or not to change a multiple choice answer. *U.S. Army Med Department J*. 2014;86.
13. Mueller DJ, Shwedel A. Some correlates of net gain resultant from answer changing on objective achievement test items. *J Educ Meas*. 1975;12(4):251-254.
14. Jordan L, Johnson D. The relationship between changing answers and performance on multiple-choice nursing examinations. *J Nurs Educ*. 1990 Oct;29(8):337.
15. Ballance C. Students' expectations and their answer-changing behavior. *Psychol Rep*. 1977;41(1):163-166.
16. Bauer D, Kopp V, Fischer MR. Answer changing in multiple choice assessment change that answer when in doubt—and spread the word. *BMC Med Educ*. 2007;7(1):28.
17. Cox-Davenport R, Haynes P, Lawson T. A mixed-methods approach to evaluating student nurses changing answers on multiple choice exams. *Hear J*. 2014 Apr;67(4):6.

18. Rangel RH, Möller L, Sitter H, Stibane T, Strzelezyk A. Sure, or unsure? Measuring students' confidence and the potential impact on patient safety in multiple-choice questions. *Med Teach*. 2017;39(11):1189-1194.
19. McMorris RF, DeMers LP, Schwarz SP. Attitudes, behaviors, and reasons for changing responses following answer-changing instruction. *J Educ Meas*. 1987;24(2):131-143.
20. Ferguson KJ, Kreiter CD, Peterson MW, Rowat JA, Elliott ST. Is that your final answer? Relationship of changed answers to overall performance on a computer-based medical school course examination. *Teach Learn Med*. 2002;14(1):20-23.
21. Americans with Disabilities Act. n.d. Retrieved January 30, 2023, from <https://www.ada.gov/>.
22. Nieswiadomy RM, Arnold WK, Garza C. Changing answers on multiple-choice examinations taken by baccalaureate nursing students. *J Nurs Educ*. 2001 Mar;40(3):142-144.
23. Stoffer G, Davis K, Brown J. The consequences of changing initial answers on objective tests: a stable effect and a stable misconception. *J Educ Res*. 1977;70(5):272-277.
24. Harvill L, Davis G. Medical students' reasons for changing answers on multiple-choice tests. *Acad Med*. 1997 /10/01;72(10):97.
25. McConnell MM, Regehr G, Wood TJ, Eva KW. Self-monitoring and its relationship to medical knowledge. *Adv Health Sci Educ*. 2012;17(3):311-323.
26. McMorris RF, Weideman AH. Answer changing after instruction on answer changing. *Meas Eval Couns Dev*. 1986;19(2):93-101.
27. Vidler D, Hansen R. Answer changing on multiple-choice tests. *J Exp Educ*. 1980;49(1):18-20.
28. Pascale PJ. Changing initial answers on multiple-choice achievement tests. *Meas Eval Guidance*. 1974;6(4):236-238.
29. 2023 Best Medical Schools. n.d. US News & World Report. Retrieved January 31, 2023, from <https://www.usnews.com/best-graduate-schools/top-medical-schools/research-rankings>.
30. Westin J. 2019, Nov. 26. Medical School Tiers and the Scores You Need. Retrieved January 31, 2023, from <https://jackwestin.com/resources/blog/medical-schools-tiers-and-the-scores-you-need>.
31. Kao G, Thompson JS. Racial and ethnic stratification in educational achievement and attainment. *Annu Rev Sociol*. 2003;29(1):417-442.
32. Chung E, Turnbull D, Chur-Hansen A. Who are non-traditional students? A systematic review of published definitions in research on mental health of tertiary students. *Educ Res Rev*. 2014;9(22):1224-1238.