

A Preliminary Study of the Value of Personality Assessment in Medical School Admissions within the United States

Peter Eveland

California University of Science and Medicine

Sabrina Wilhelm

California University of Science and Medicine

Stephanie Wong

California University of Science and Medicine

Lissett G. Prado

California University of Science and Medicine

Sanford H. Barsky (✉ BarskyS@cusm.org)

California University of Science and Medicine

Research Article

Keywords: Medical school admissions, medical school interviews, personality assessments, Neo Personality Inventory-Revised Test

Posted Date: March 16th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1429390/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Allopathic medicine faces a daunting challenge of selecting the best applicants because of the very high applicant / matriculant ratio. The quality of graduates ultimately reflects the quality of medical practice. Alarming recent trends in physician burnout, misconduct and suicide raise questions of whether we are selecting the right candidates. The United States lags far behind the United Kingdom and Europe in the study of non-cognitive tests in medical school admissions. Although more recently, medical schools in both the United Kingdom, Europe and the United States have begun to use situational judgement tests such as CaSPER and the SJT, recently developed by the AAMC and that these tests are, in a sense non-cognitive in nature, direct personality tests per se have not been utilized. Although personality is one indelible component of the human condition, we have historically used, in the admissions process within the US, knowledge, reasoning and exam performance, all of which can be improved with practice.

Methods: A popular personality measurement used over the past two decades within the US in business and industry, but not medical school has been the NEO-PI-R Test. This test has not been utilized in allopathic medicine probably because of the paucity of exploratory retrospective and validating prospective studies. The hypothesis which we tested was whether NEO-PI-R traits exhibit consistency between two institutions and whether their values show promise in predicting academic performance.

Results: Our retrospective findings indicated both interinstitutional consistencies and both positive and negative predictive values for certain traits whose correlative strengths exceeded traditional premed metrics (MCAT, GPA, etc.) for early academic performance.

Conclusions: Our exploratory studies should catalyze larger and more detailed confirmatory studies designed to validate the importance of personality traits not only in predicting early medical school performance but also later performance in one's overall medical career.

Background

Intelligence and personality are two indelible components of the human condition. Cognitive skills, knowledge, reasoning and exam performance, on the other hand, can be acquired and improved through practice.¹ We nearly exclusively use the latter in the medical school admission process in the United States, and largely ignore personality, at least by formal assessment. Alarming recent trends in physician burnout, misconduct and suicide raise additional questions of whether we are selecting the right candidates in our medical school admissions process. It is not entirely clear why in the United States we continue to use only premed cognitive assessments in selecting whom to accept to medical school. Not only have we continually ignored non-cognitive assessments in the admissions process but we have not even conducted retrospective or prospective studies examining their potential value in predicting early medical school performance or later performance in one's overall medical career. This dearth of US studies stands in contrast to those in the United Kingdom and Europe which consist of a number of large

cohort studies examining non-cognitive testing which include both modifiable as well as non-modifiable personality traits and their predictive value during and at completion of medical school.²⁻¹¹

Although more recently, medical schools in both the United Kingdom, Europe and the United States have begun to use situational judgement tests such as CaSPER and the SJT, recently developed by the AAMC and that these tests are, in a sense non-cognitive in nature,^{12,13} direct personality tests per se have not been utilized. An increasingly popular formal measurement of personality, however, which has evolved over the past two decades, is the Neo Personality Inventory – Revised (NEO-PI-R) Test, a measurement of five major domains of personality as well as six facets that define each of the domains (Table 1).^{14,15} The NEO-PI-R is a psychological personality inventory consisting of the Five Factor Domain (Model): Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness to Experience. The test also measures six subordinate dimensions, known as ‘facets’ of each of the five factor model personality domains. The NEO-PI-R consists of 240 items of descriptions of behavior answered on a five point scale, ranging from “strongly disagree” to strongly agree”.¹⁴ The test is available both online and in paper form and has been used widely in the evaluation of employee applications in business, industry, law enforcement and selectively high pressured occupations, e.g., air traffic controllers.^{16,17} The test has not formally or officially been used in the allopathic medical school admissions process in the United States for reasons that are not totally clear. Perhaps one reason it has not been used is that there have been a paucity of exploratory retrospective and validating prospective studies examining the value of formal personality assessment in the medical school setting.

It can be argued however that since the medical school admissions process uses either in-person or virtual interviews, that some aspects of applicant personality invariably surface during the interview process and may influence decisions of acceptance.¹⁸⁻²⁰ However that is different than a formal, systematic, objective, quantitative and reproducible measurement of personality as can be offered by the NEO-PI-R test. Overall, there has been, in fact, only a paucity of studies examining personality traits of medical applicants and matriculants.²¹ Exploratory retrospective and confirmatory prospective studies of the NEO-PI-R are first needed to justify its routine use in the medical school admissions process. For these studies to be valid, the NEO-PI-R test must be separately administered to all applicants granted an interview but must not at all be used, at least initially, to influence the admissions process and interviewees must not be told whether the test will influence or not influence the decision process. This is exactly what we did in the present study. The hypothesis which we tested was whether personality traits measured by NEO-PI-R are consistent between two institutions and whether they have value in predicting academic success as well as failure, greater than traditional premed metrics (MCAT, GPA, etc).

Methods

This study was conducted under strict FERPA guidelines. All data had been collected as part of the routine admissions process and subjects de-identified. The present study was approved by CUSM’s IRB (HS-2020-04). We had previously collected 2 year’s worth of matriculant data from Mercer University

School of Medicine (MUSM) under an approved IRB (H0312123). All raw data analyzed in the study is provided as Supplementary Information designated as Raw Data - Appendix 1–4.

This study was conducted blindly. The individuals at both institutions who administered the NEO-PI-R to the interviewees and recorded the results did not participate in any other aspects of the medical school interview or admissions process, did not interact with members of the Admissions Committee in any way nor participate in the deliberations or decisions of the Admissions Committee. Numerical values of NEO traits and subtraits from all interviewees from the classes of 2022 and 2023 at CUSM and all students from the classes of 2006 and 2007 at MUSM were descriptively summarized using means, standard deviations, minimums, maximums, ranges, and variances. Comparisons of means between NEO personality traits of CUSM and MUSM students were conducted using one-way ANOVA to determine any statistically significant differences. An alpha value of 0.05 was considered significant.

NEO traits between those with good vs poor performance at MUSM were compared using an independent sample t-test. Poor performance had three subcategories: repeating a single course, repeating multiple courses or dropping out of school. NEO traits of students with good performance were compared to NEO traits of students with poor performance.

Comparisons between accepted students and rejected students at CUSM were assessed by conducting an independent sample t-test for NEO traits of accepted students and NEO traits of rejected students for both classes of 2022 and 2023, and each year individually.

NEO traits of accepted and rejected students from CUSM and MUSM were subsequently compared using independent sample t-tests in the following categories: (1) MUSM vs CUSM All Accepted, (2) MUSM vs CUSM All Rejected, (3) MUSM vs CUSM Year 1 Accepted, (4) MUSM vs CUSM Year 1 Rejected, (5) MUSM vs CUSM Year 2 Accepted, (6) MUSM vs CUSM Year 2 Rejected.

Correlations between different NEO traits in CUSM students were calculated using a 2-tailed Pearson bivariate correlation and charted as a matrix. An alpha value of 0.05 was considered significant. Correlations between NEO traits in CUSM students and class rank were similarly calculated. Correlations between NEO traits in CUSM students and typical premedical admissions metrics (MCAT, CGPA, BCPM) as well as medical school performance metrics (average NBME, MCQ, LAB, CP, IRAT, OSCE, and CRS).

Differences in NEO traits between male and female accepted and rejected applicants were compared using independent sample t-tests for both CUSM classes of 2022 and 2023.

A more detailed enumeration of the tests and comparisons that were conducted is provided (Table 2).

Results

The hypothesis which we tested was whether personality traits as measured by the NEO-PI-R Test have predictive value in early medical school performance and whether this predictive value was stronger than traditional premed metrics (MCAT, GPA, etc.). Obviously, if support for this hypothesis could be obtained

from this study, it would argue possibly for an expanded role of the NEO-PI-R Test in the medical school admissions process or at least for additional confirmatory retrospective and validatory prospective studies.

At MUSM, the Admissions Committee did not formally use the NEO-PI-R test to evaluate prospective applicants and were completely blinded to the NEO-PI-R Test results. Therefore, any correlations between personality scores and academic performance were made on an unselected and therefore unbiased population. In the present study we re-analyzed the MUSM raw data. We also made comparisons between the MUSM and CUSM data.

The present study also examined 2 years of CUSM applicant and matriculant data for NEO-PI-R, premedical parameters, demographic data and medical school performance data for potential predictive value of the NEO-PI-R v traditional premed parameters.

Even though the MUSM data and the CUSM data were derived from different populations of medical school applicants, approximately 15 years apart, with different demographic features, (eg.,the male / female ratio was much higher at MUSM), from different schools with different admission criteria, and from different geographic areas of the United States, the NEO-PI-R was remarkably consistent in the personality mean scores and ranges between the two groups of students. 29 of 30 facets of personality showed no differences in score distribution between the populations ($p = .87$; $p = .78$). The single facet showing a difference between the two populations was (A6) Tender-Mindedness ($p = .007$). This facet accounted for a difference in its member domain (A) Agreeableness ($p = .034$). The fact that 29/30 personality facets showed no differences between the MUSM and CUSM student populations demonstrated the remarkable consistency of the NEO-PI-R. This consistency spanned decades, schools, demographics and geographies.

Re-analysis of the MUSM data revealed a number of interesting findings. For one there were significant differences in one major personality domain as well as many of its facets between males v females. The one major domain which showed differences was (C) Conscientiousness with females scoring higher ($p = .012$). Females also scored higher in two of its facets: (C2) Order ($p = .026$) and (C6) Deliberation ($p = .02$). Within the domain of (A) Agreeableness, the facet (A4) Compliance showed higher scores in males ($p = .032$).

A number of personality domains and facets correlated with either academic success or failure in both males and females. Academic success was defined by separate and cumulative course performance and academic failure was defined as having to repeat a single course or multiple courses or dropping out of school. The predictive values of these personality domains and facets were compared to the predictive values of traditional premed metrics like MCAT VR (verbal reasoning), MCAT PS (physical sciences: chemistry, physics) and MCAT BS (biological sciences: biology, biochemistry, genetics, physiology, molecular biology, microbiology, evolution, organic chemistry). At the time of the MUSM study, the MCAT was divided into MCAT VR, MCAT PS and MCAT BS. The MCAT BS scores positively correlated with 7 different courses performances ($p = .05$; Pearson 0.6) and the MCAT PS positively correlated with 2

course performances ($p = .05$; Pearson 0.6) whereas MCAT VR negatively correlated with 4 course performances ($p = .01$; Pearson -0.7). However, none of the MCAT scores correlated with academic failure.

A number of personality domains and facets also correlated positively (significant and positive Pearson coefficients) with course performances. Most of these fell within the (C) Conscientiousness domain which included (C3) Dutifulness, 5 courses ($p = .03$; Pearson -0.8); (C4) Achievement Striving, 4 courses ($p = .04$; Pearson 0.7); and (C5) Self-Discipline, 7 courses ($p = .02$; Pearson 0.9). Collectively, the facets within the (C) Conscientiousness domain correlated with academic success in more courses than the MCAT BS and MCAT PS scores.

However, the most striking finding in the MUSM data was the negative correlations (significant and negative Pearson coefficients) with academic failure. The personality domains and facets which provided strong negative correlations with academic failure (repeating a single course, multiple courses or dropping out of school) fell mainly within the (N) Neuroticism domain including facets (N2) Angry Hostility ($p = .05$; Pearson -0.7) and (N3) Depression ($p = .05$; Pearson -0.7) and the (O) Openness to Experience domain which included facets (O2) Aesthetics ($p = .05$; Pearson -0.7) and (O3) Feelings ($p = .05$; Pearson -0.7). The facets within the (O) Openness to Experience domain negatively correlated with repeating not just one but multiple courses ($p = .028$; Pearson -0.9). Select personality domains and facets therefore potentially add value to the admissions process as a negative predictor of academic failure.

Similarly to the MUSM students whose admissions to medical school were not at all based on the NEO-PI-R test, CUSM did not use the NEO-PI-R test to formally influence admissions. In the first class which was admitted (the class of 2022), 29 of 30 facets of personality predictably showed no differences in score distribution between the accepted vs rejected applicants ($p = .250$). In the second class which was admitted (the class of 2023), there were differences in only 1 domain: (N) Neuroticism. In fact, all of the facets within this domain showed differences between accepted vs. rejected applicants ($p = .02$). Although the NEO-PI-R test was not formally used as an Admissions Criteria and whose results were not made available to the Admissions Committee, it was entirely possible that the interviewers were sensitive to neurotic personality traits of certain applicants that negatively impacted their decisions on acceptance. It would seem then from this observation that this domain may have factored into the admission decision.

Analysis of the CUSM data revealed both similarities and differences compared to the MUSM data. The personality profiles of males vs females were again different but mainly fell in facets within the (E) Extraversion, (O) Openness to Experience and (A) Agreeableness domains ($p = .02$). CUSM accepted approximately equal number of males and female students whereas MUSM accepted only a limited number of female students at that time. The difference in male / female ratio between the two classes could explain the discrepancy in the differing personality facets.

Since there is currently more of an emphasis on evaluating medical school student performance to comply with the rigors of the LCME accreditation process than there was 15 years ago, CUSM used a number of performance metrics that were not available at MUSM which included Multiple Choice Questions (MCQs), NBME (both raw and scaled), Laboratory, Case Presentation, iRAT, OSCE, Course Final Grade (derived from a composite of measurements depicted below) and Overall Averages.

The Course Final Grade (Raw Score) was derived from a composite of the detailed measurements as depicted (Table 3). In addition, other premedical metrics that were available included overall MCAT, overall GPA (grade point average) and BCPM (Biology, Chemistry, Physics, Mathematics) grade point average. Presently, only an overall- MCAT score was available because the MCAT was no longer broken into MCAT VR, MCAT PS and MCAT BS as it was for the MUSM data.

At CUSM, presently, academic failure was defined as the need to repeat a course but since no CUSM students to date, however, have been required to repeat a course due to students' 100% successful attempts at remediation, academic failure per se could not be correlated with NEO-PI-R measurements, Academic performance (success or lack thereof) based on various assessments including class rank (Table 3) could be measured and was used in this study.

With traditional premed metrics, MCAT scores surprisingly did not significantly correlate with any of the above-mentioned assessments ($p = .5$). However, BCPM significantly correlated with 3 of the assessments ($p = .01$; Pearson 0.7) and was therefore the best of the objective metrics.

However, the most striking finding discovered was the very strong negative correlations (significant and negative Pearson coefficients) with academic performance by certain personality domains and facets. The personality domains and facets which provided strong negative correlations with academic performance fell mainly within the (N) Neuroticism domain including facets (N2) Angry Hostility, (N3) Depression, (N5) Impulsiveness and (N6) Vulnerability (all, $p = .02$; Pearson -0.8). These facets negatively correlated with as many as 4 of the assessments, which were more assessments than those that correlated with the BCPM. Interestingly, the (N) Neuroticism domain including facets (N2) Anger Hostility and (N3) Depression were also the same personality domain and facets that predicted academic failure at MUSM.

Discussion

Allopathic medical schools continue to receive many more applications than class openings and therefore have an opportunity to select the "right" and "best" applicants. However the recently increasing rates of physician burnout, professional misconduct and physician suicide all raise questions as to whether we are selecting the right applicants. It is certainly possible and even plausible that non-cognitive assessments of such things as personality traits could provide potential input in the selection of candidates to decrease these negative outcomes of long term practice. Historically applicants in the United States have been selected on the basis of fairly standard premedical metrics which include GPA, selected science and math GPA and MCAT scores. These metrics produce a fairly homogeneous pool of

selected applicants. Yet medical school applicants are heterogeneous in terms of interests, motivations, career goals and personality traits. Personality represents a component of the human condition which has not been adequately explored in the medical school admission process nor adequately used to predict future career success or failure in medicine.

Certainly it could be argued that students who aspire to a career in family medicine to treat the underserved more likely possess different personality traits than aspiring physician-scientists who are willing to forgo the practice of the art of medicine in favor of its science. Yet probably both categories of students exhibit a similar range of traditional premed metrics like GPA and MCAT scores that serve as the gateway to their admission.

Although there have been a number of studies in the United States that have examined personality traits of medical students, there have been few studies that have examined these traits as predictors of medical school performance.¹⁸⁻²¹ And certainly there have been no studies that have examined personality factors as predictors of ultimate career success or failure. Furthermore, we are not aware of any allopathic medical school in the United States that formally uses scored personality assessments such as those of the NEO-PI-R test as a criterion in determining admission.

The United States therefore lags far behind the United Kingdom and Europe in the study and use of non-cognitive tests in medical school admissions in predicting subsequent performance. It is fair to say that the US is in its infancy with regards to non-cognitive testing. The reasons for this are not entirely clear. Numerous studies in the United Kingdom, Europe and other non-US countries have investigated the role and importance of non-cognitive tests in medical school admissions and their role in predicting medical school performance.^{10, 22-32} These studies used four types of non-cognitive tests including libertariancommunitarian, narcissism, aloofness, confidence and empathy (NACE), self-esteem, optimism, control, self-discipline, emotional-nondefensiveness (END) and faking (MEARS) and combinations thereof.¹⁰ Performance measurements included the Educational Performance Measure (EPM) and the exit situational judgement test (SJT). Multilevel regression analyses showed that END predicted EPM and SJT and that two facets of NACE, aloofness and empathy predicted SJT. Although these studies showed some significant correlations they exhibited overall low effect sizes and an inconsistent picture. These personality tests consisted of a very broad range of characteristics which could be separated into so-called modifiable traits such as social and communication skills, perseverance, resilience and motivation and so-called non-modifiable traits such as neuroticism and extraversion.

These studies specifically did not use the NEO-PI-R Test which measures the so-called "Big Five": Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness to Experience. It should again be emphasized that the NEO-PI-R Test measures non-modifiable or indelible and stable aspects of personality whereas NACE and MEARS are thought to measure, at least, in part modifiable traits. Measuring modifiable traits bring to any study a type of confounding which is difficult to control for. It is interesting that the one sole study conducted in Europe that did use only the "Big Five" showed that certain traits did correlate with academic performance.⁴

In the vast majority of these non-US studies, it was made clear to the candidates that that the non-cognitive tests would not be used as a basis for admissions and so it could be argued that the candidates were less motivated to take the test seriously. Furthermore none of these studies measured long term outcomes of medical performance.

In order to make a case that formal personality assessment has a role in the Admissions process in the United States, we first needed to show in our study that formal quantitative personality assessment correlated with medical school performance and that this correlation was observed on an unselected and therefore unbiased population. In both the MUSM and CUSM classes, this opportunity presented itself.

In our study we conducted a large number of blind analyses without any preconceived rationale because we did not want to bias our results. It can be argued that we approached this study largely as a fishing expedition. However this “fishing” approach was appropriate and justified given the dearth of previous studies on the utility of the NEO-PI-R in medical school admissions. Our results which not only show statistical significance but strong Pearson correlations in the setting of a relatively small sample and our demonstrations of stronger performance correlations of select NEO traits v standard premed metrics also argues against a type 1 error and suggest that our preliminary studies be followed up with larger confirmatory retrospective studies and eventual validatory prospective studies.

Given that CUSM has not yet even graduated a class, the true predictive value of the personality test can not yet be fully evaluated and therefore this study must be considered preliminary. In particular due to the relatively small numbers, we were only able to conduct bivariate analyses of the different personality traits and academic success. Since there are other well known predictors of academic success such as MCAT scores, that could colinearly distribute with one or more of the personality test scores, it would be important once more data is available to establish that personality scores in a multivariate model are superior or at least show that the cognitive values do not differ significantly between students with different outcomes on the personality test. Similarly although we noted that there was a difference between some of the personality values between males and females, due to the limited data which was available to us, we did not adjust for this possible confounding variable in other comparisons.

Furthermore with the growing popularity of the non-cognitive situational judgement tests such as CaSPER and the SJT while different than personality tests per se, it would be equally important to directly compare direct personality tests with these non-cognitive tests to determine whether personality tests have better predictive value of medical school performance. An expanded data set would allow these additional comparisons.

In any correlative or experimental study of medical education such as this one, it is important to provide the conceptual framework which serves as background. Conceptual frameworks represent ways of thinking about a problem or study.³³ Conceptual frameworks can come from theories, models or best practices but all of these can be challenged as myths, if the evidence suggests the contrary.³⁴ Historically it has been assumed that measurements of cognitive skills, learning, knowledge, reasoning and exam

performance, all of which can be improved through practice, are the best predictors of not only medical school success but overall career success in medicine. However these assumptions may prove faulty as personality, an indelible component of the human condition, may ultimately be more important in predicting both medical school performance as well as overall career success or failure. Furthermore it can be reasoned that if we can measure and delineate personality, we might be able to tailor individual instruction to selectively nurture individuals with certain personality traits and, in a sense, develop a form of personalized instruction. If we can achieve both, then without question, personality assessment should be used as a gateway, at least in part, to admission.

Conclusions

Our retrospective exploratory analyses of the data at MUSM and CUSM argue for the importance of measuring personality domains and facets provided by the NEO-PI-R to provide prognostic information on academic performance. Obviously, these initial findings must be evaluated both in subsequent classes and in the present classes when more performance data, e.g. USMLE scores and clinical performance become available. Our retrospective analyses should be subsequently examined with both confirmatory prospective studies and future long term validation studies that examine not only medical school performance but overall career performance. These studies would fulfill the often neglected LCME mandate that medical schools in the United States and Canada select applicants who possess the intelligence, integrity *and personal and emotional characteristics* necessary to become competent physicians in the practice of medicine.

Abbreviations

CUSM, California University of Science and Medicine; MUSM, Mercer University School of Medicine; IRB, institutional review board; FERPA, Family Educational Rights and Privacy Act; LCME, Liaison Committee on Medical Education; ANOVA, analysis of variance; USMLE, United States Medical Licensing Examination; MCAT, Medical college admission test; GPA, grade point average

Declarations

Ethics Approval and Consent to Participate

This study was conducted under FERPA guidelines. All data had been collected as part of the routine admissions process and subjects de-identified. The present study was approved by the California University of Science and Medicine's Institutional Review Board (HS-2020-04) who granted a waiver of consent. We had previously collected 2 year's worth of matriculant data from Mercer University School of Medicine (MUSM) under Mercer University School of Medicine's Institutional Review Board (H0312123) who also granted a waiver of consent.

Consent for Publication

Not applicable

Availability of Data and Materials

All raw data, in de-identified format, analyzed in this study are included in this published article in its Supplementary Information designated as Raw Data - Appendix 1-4.

Competing Interests

The authors declare that they, at the present time, have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. None of the sources of support listed influenced the collection, analysis and interpretation of data, the generation of the hypothesis, the writing of the manuscript or the decision to submit the manuscript for publication.

Funding

The work was supported by the California University of Science and Medicine and the Dr. Carolyn S. Glaubenslee Endowed Cancer Center Directorship.

Authors' Contributions

All authors made intellectual contributions to the work and have written portions of the manuscript. A Peter Eveland originated the hypothesis that the NEO test was predictive of medical school performance. Sabrina R. Wilhelm administered the exam to all CUSM applicants and matriculants and de-identified all subjects. Stephanie Wong conducted detailed statistical analyses of the data blindly. Lissett G. Prado assisted with the analysis of all of the data. Sanford H Barsky designed the overall approach of the study to demonstrate the possible stronger predictive values of select personality traits over traditional premed metrics.

Acknowledgements

The authors wish to thank Jason Crowley and Clarissa Barra of CUSM's Office of Medical Education, CUSM's Instructional and Informational Technology Services for enabling videoconferencing coauthor communications and Louise Borda for providing operational assistance in the creation of this manuscript.

Authors' Information

A Pete Eveland is Senior Associate Dean of Student Affairs and Admissions. **Sabrina R. Wilhelm** is Director of Academic Skills and Career Advising. **Stephanie Wong** is a third year medical school student. **Lissett G. Prado** is Executive Assistant to the Associate Dean of Students. **Sanford H. Barsky** is

References

1. Shen H, Comrey AL. Predicting medical students' academic performances by their cognitive abilities and personality characteristics. *Acad Med*. 1997;72(9):781–786. doi:10.1097/00001888-199709000-00013
2. Adams, J., Bore, M., Childs, R., Dunn, J., McKendree, J., Munro, D., et al. (2015). Predictors of professional behaviour and academic outcomes in a UK medical school: A longitudinal cohort study. *Medical Teacher*, 2015(37), 868–880.
3. Adams, J., Bore, M., McKendree, J., Munro, D., & Powis, D. (2012). Can personal attributes of medical students predict in-course examination success and professional behaviour? An exploratory prospective cohort study. *BMC Medical Education*, 12:69. <http://www.biomedcentral.com/1472-6920/12/69/abstract>.
4. Lievens, F., Coetsier, P., De Fruyt, F., & De Maeseneer, J. (2002). Medical students' personality characteristics and academic performance: A five factor model perspective. *Medical Education*, 36(11), 1050–1056.
5. Dowell, J., Lumsden, M. A., Powis, D., Munro, D., Bore, M., Makubate, B., & Kumwenda, B. (2011). Predictive validity of the personal attributes assessment for selection of medical students in Scotland. *Medical Teacher* 33, e485–e488
<http://informahealthcare.com/doi/abs/10.3109/0142159X.2011.599448?prevSearch=allfield%253A%2528jon%2Bdowell%2529&searchHistoryKey>.
6. Patterson, F., Carr, V., Zibarras, L., Burr, B., Berkin, L., Plint, S., et al. (2009). New machine-marked tests for selection into core medical training: Evidence from two validation studies. *Clinical Medicine*, 9, 417–420.
7. McManus, I. C., Woolf, K., Dacre, J., Paice, E., & Dewberry, C. (2013). The academic backbone: Longitudinal continuities in educational achievement from secondary school and medical school to MRCP(UK) and the specialist register in UK medical students and doctors. *BMC Medicine* 11(1), 242. <http://www.biomedcentral.com/content/pdf/1741-7015-11-242.pdf>.
8. Patterson, F., & Ashworth, V. (2011). Situational judgement tests; the future for medical selection? *British Medical Journal*. <http://careers.bmj.com/careers/advice/view-article.html?id=20005183>
9. Schuwirth, L., & Cantillon, P. (2005). The need for outcome measures in medical education. *British Medical Journal*, 331, 977.
10. MacKenzie, R.K., Dowell, J., Ayansina, D., Cleland, J.A. (2017). Do personality traits assessed on medical school admission predict exit performance? A UK-wide longitudinal cohort study. *Adv in Health Sci Educ* 22, 365–385.

11. Patterson, F. (2013). Selection for medical education and training: Research, theory and practice. In K. Walsh (Ed.), *Oxford Textbook for Medical Education* (pp. 385–397). Oxford: Oxford University Press.
12. Tiffin PA, Paton LW, O'Mara D, MacCann C, Lang JWB, Lievens F. Situational judgement tests for selection: Traditional vs construct-driven approaches. *Med Educ*. 2020 Feb;54(2):105–115. doi: 10.1111/medu.14011. Epub 2019 Dec 23. PMID: 31872483.
13. de Visser M, Fluit C, Cohen-Schotanus J, Laan R. The effects of a non-cognitive versus cognitive admission procedure within cohorts in one medical school. *Adv Health Sci Educ Theory Pract*. 2018 Mar;23(1):187–200. doi: 10.1007/s10459-017-9782-1. Epub 2017 Jun 10. PMID: 28601913; PMCID: PMC5801390.
14. McCrae RR, Costa PT Jr. Validation of the five-factor model of personality across instruments and observers. *J Pers Soc Psychol*. 1987;52(1):81–90. doi:10.1037//0022-3514.52.1.81
15. Briggs SR. Assessing the five-factor model of personality description. *J Pers*. 1992;60(2):253–293. doi:10.1111/j.1467-6494.1992.tb00974.x
16. Costa PT Jr, McCrae RR. Stability and change in personality assessment: the revised NEO Personality Inventory in the year 2000. *J Pers Assess*. 1997;68(1):86–94. doi:10.1207/s15327752jpa6801_7
17. Costa PT, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: robust and surprising findings. *J Pers Soc Psychol*. 2001;81(2):322–331. doi:10.1037/0022-3514.81.2.322
18. Schripsema NR, van Trigt AM, van der Wal MA, Cohen-Schotanus J. How Different Medical School Selection Processes Call upon Different Personality Characteristics. *PLoS One*. 2016;11(3):e0150645. Published 2016 Mar 9. doi:10.1371/journal.pone.0150645
19. Kulasegaram K, Reiter HI, Wiesner W, Hackett RD, Norman GR. Non-association between Neo-5 personality tests and multiple mini-interview. *Adv Health Sci Educ Theory Pract*. 2010;15(3):415–423. doi:10.1007/s10459-009-9209-x
20. Obst KU, Brühem L, Westermann J, Katalinic A, Kötter T. Are the results of questionnaires measuring non-cognitive characteristics during the selection procedure for medical school application biased by social desirability?. *GMS J Med Educ*. 2016;33(5):Doc75. Published 2016 Nov 15. doi:10.3205/zma001074
21. Stratton TD, Elam CL. A holistic review of the medical school admission process: examining correlates of academic underperformance. *Med Educ Online*. 2014;19:22919. Published 2014 Apr 1. doi:10.3402/meo.v19.22919
22. Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, 37, 426–441.
23. Lumsden, M. A., Bore, M., Millar, K., Jack, R., & Powis, D. (2005). Assessment of personal attributes in relation to admission to medical school. *Medical Education*, 39, 258–265.
24. Manuel, R. S., Borges, N. J., & Gerzina, H. A. (2005). Personality and clinical skills: Any correlation? *Academic Medicine*, 80, S30–S33.

25. Munro, D., Bore, M. R., & Powis, D. A. (2005). Personality factors in professional ethical behaviour: Studies of empathy and narcissism'. *Australian Journal of Psychology*, 57, 49–60.
26. Munro, M., Bore, M., & Powis, D. (2008). Personality determinants of success in medical school and beyond: "steady, sane and nice". In S. Boag (Ed.), *Personality down under perspectives from Australia* (pp. 103–112). New York: Nova Science Publishers Inc.
27. Norman, G. (2015). Identifying the bad apples. *Advances in Health Sciences Education*, 20, 299–303.
28. O'Connell, M. S., Hartman, N. S., McDaniel, M. A., Grubb, W. L., & Lawrence, A. (2007). Incremental validity of situational judgement tests for task and contextual job performance. *International Journal of Selection and Assessment*, 15, 19–29.
29. Powis, D. A., Bore, M. R., Munro, D., & Lumsden, M. A. (2005). Development of the personal qualities assessment as a tool for selecting medical students. *Journal of Adult & Continuing Education*, 11, 3–14
30. Patterson, F., Knight, A., Dowell, J., Nicholson, S., & Cleland, J. A. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education*, 50(1), 36–60.
31. Prideaux, D., Roberts, C., Eva, K., Centeno, A., McCrorie, P., McManus, C., et al. (2011). Assessment for selection for the health care professions and specialty training: International consensus statement and recommendations. *Medical Teacher*, 33, 215–223.
32. Roberts, C., Walton, M., Rothnie, I., Crossley, J., Lyon, P., Kumar, K., et al. (2008). Factors affecting the utility of the multiple mini-interview in selecting candidates for graduate-entry medical school. *Medical Education*, 42, 396–404.
33. Bordage G. Conceptual frameworks to illuminate and magnify. *Med Educ*. 2009;43(4):312–319. doi:10.1111/j.1365-2923.2009.03295.
34. Martimianakis MAT, Tilburt J, Michalec B, Hafferty FW. Myths and social structure: The unbearable necessity of mythology in medical education. *Med Educ*. 2020;54(1):15–21. doi:10.1111/medu.13828

Tables

Tables 1 to 3 are available in the Supplemental Files section.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.docx](#)
- [Table2.docx](#)
- [Table3.docx](#)
- [RawDataAppendix1.xlsx](#)

- [RawDataAppendix2.xlsx](#)
- [RawDataAppendix3.xlsx](#)
- [RawDataAppendix4.xlsx](#)