

Aptitude, achievement and competence in medicine: a latent variable path model

V. Terri Collin · Claudio Violato · Kent Hecker

Received: 12 October 2007 / Accepted: 24 April 2008 / Published online: 15 May 2008
© Springer Science+Business Media B.V. 2008

Abstract To develop and test a latent variable path model of general achievement, aptitude for medicine and competence in medicine employing data from the Medical College Admission Test (MCAT), pre-medical undergraduate grade point average (UGPA) and demographic characteristics for competence in pre-clinical and measures of competence (United States Licensure Examination {USMLE} Steps 1, 2, and 3). Data were gathered on 839,710 participants from 1991 to 2000 on demographic and school variables, UGPA, MCAT subtest scores and Steps 1, 2, and 3 of the United States Licensure Examination (USMLE). However, subsets of the total 839,710 participants included in the database were used for various analyses and the testing of a latent variable path model (LVPA). A number of preliminary descriptive and inferential techniques were used to confirm previous hypotheses and stated relationships amongst the variables of interest to the present study. Through development and testing of a latent variable path model, three latent variables measured by UGPA (*general achievement*), subscales of the MCAT (*aptitude for medicine*), and Steps 1, 2, and 3 of the USMLE (*competence in medicine*) were identified which resulted in a comparative fit index = .932 of the model to a large sample ($n = 20,714$). In a confirmatory latent variable path model we were able to identify theoretical constructs, *aptitude for medicine*, *general achievement*, and *competence in medicine* and their interrelationships. These are distinct but interrelated latent variables.

Keywords Medical school performance · MCAT · Predictive validity · Latent variable path analyses

V. T. Collin (✉)

Department of Surgery, University of Pittsburgh Medical Center,
200 Lothrop St. Room F677PUH, Pittsburgh, PA 15213, USA
e-mail: collinvl@upmc.edu

C. Violato

Medical Education and Research Unit, Faculty of Medicine, University of Calgary,
Calgary, AB, Canada

K. Hecker

Faculty of Veterinary Medicine, University of Calgary, Calgary, AB, Canada

The predictive validity of the Medical College Admission Test (MCAT) for criterion measures such as performance in medical school and on licensing examinations such as the United States Medical Licensure Examinations (USMLE Steps 1, 2, and 3), continues to be an important issue notwithstanding the substantial research that has been conducted (Albanese et al. 2003; Basco et al. 2002; Jones and Thomae-Forgues 1984; Donnon et al. 2007). The MCAT is intended to assess the ability to acquire knowledge in medical school as well as higher-order processes such as clinical reasoning and the application of knowledge into clinical practice (*aptitude for medicine*). Accordingly, the MCAT is presumed to not only measure biological and physical sciences knowledge, but also to assess higher order cognitive processes such as verbal reasoning and critical thinking.

Numerous questions have arisen, however, about the predictive and construct validity of the MCAT as a measure of *aptitude for medicine* and its role in the future development of achievement among medical professionals (Donnon and Violato 2006; Violato and Donnon 2005). The focus of the present study, therefore, was to investigate further the predictive and other criterion-related validity of MCAT scores, and adduce evidence for the construct validity of the MCAT. Accordingly, we wished to develop and test a latent variable path model that not only identifies important theoretical constructs, but also explicates the structural paths interconnecting the latent variables.

Introduction

Measuring professional aptitude and academic achievement have been large-scale activities for identifying and selecting those persons most suited for medical school. Although there have been and continue to be challenges (theoretical and psychometric) to finding a “best-fit” model for medical school selection, the use of the MCAT and other cognitive based pre-admission variables (i.e., pre-medical undergraduate grade point average—UGPA) have been the most valid measures for predicting who will succeed in medical school, particularly in the pre-clinical years.

Since the revision to the MCAT in 1991, researchers have searched for evidence of its predictive validity. Traditionally, undergraduate science and total GPA and MCAT scores (specifically biological, physical, and verbal subtests) have been generally correlated with pre-clinical competence as measured by the USMLE Step 1, and the MCAT as a single variable is considered to be the best predictor of performance of medical school courses and Steps 1, 2, and 3 of the USMLE (Donnon et al. 2007). In one study, Swanson and colleagues derived a predictive validity coefficient of $r = .52$ for MCAT total based on a large sample ($n = 11,145$) involving many medical schools (Swanson et al. 1996). This study and others have shown that the MCAT adds unique input to predicting future performance, particularly in pre-clinical criterion measures (Koenig et al. 1998). Although the MCAT has shown to be a better predictor of pre-clinical performance, it also appears to have some predictive validity of those competencies in the clerkship years that may involve higher-order cognitive processes beyond those assessed by pencil-and-paper tests. Veloski, Callahan, Xu, Hojat and Nash concluded that MCAT scores alone explained ~20% of the variance in clerkship performance and when adding UGPA the predictive power increased that much more (Veloski et al. 2000). The present study used total UGPA and MCAT subtest scores as observed variables for predicting criterion measures (USMLE Steps 1, 2, and 3) and for defining the theoretical or latent constructs assumed to underlie the process predicting the success of students before, during, and following medical school.

Overview of observed variables in study

MCAT

The chief purpose of the MCAT is to offer a reliable and valid estimate of a candidate's aptitude for mastering the goals and objectives set forth in a medical school curriculum, essentially predicting those applicants who will be successful in medical school. The MCAT is conceptualized within the literature as a preparatory indicator measuring readiness into the medical field and rigorous enough to identify those applicants who will be successful students and physicians. The current version of the MCAT (also the version used in the present study) is a multiple-choice test and was introduced in 1991. It consists of four subscales: biological sciences, physical sciences, verbal reasoning, and a writing sample. For three of the MCAT subscales (i.e., biological sciences, physical sciences and verbal reasoning), converted scores are reported and range between 1 and 15, with a mean of 8 and varying standard deviations depending on subtest (usually between 2.0 and 2.5) (Hojat et al. 2000; Association of American Medical Colleges 1991). The writing sample which consists of two-30 min essay questions, assesses analytical thinking and writing skills. The written sample is scored by a minimum of 4 raters, with a reported reliability between .68 and .78 (Gilbert et al. 2002). Unlike the numerical scores characteristic of the other three subtests, the writing sample is scored alphabetically from J to T (worst to best) (Hojat et al. 2000). The overriding response from admissions committees for using MCAT scores as a screening and selection instrument is that it provides them with a global estimate of identifying those candidates that can succeed in medical school (Mitchell et al. 1994). Moreover, as the MCAT has been found to be significantly related with medical school GPA (especially the first 2 years) as well as the three steps of the licensure exams (USMLE Steps 1, 2, and 3), which assess basic science knowledge and clinical applications, candidates with higher MCAT scores are assumed to benefit more from a medical program especially since the main objective in the first 2 years is to learn biomedical principles.

USMLE

The USMLE consists of three sections. Step 1 is administered after the second year of medical school, and assesses knowledge concerning basic science principles as well as concepts related to patient-centered care and principles of medical care (United States Medical Licensing Examination 2008). Step 2 is administered after the final year of medical school (i.e., clerkship) and assesses clinical competence namely, skills and knowledge applied in medical practice. Lastly, Step 3 is generally taken after the first year of residency and assesses problem-solving skills and one's ability to apply knowledge and fundamental principles to defined clinical scenarios (United States Medical Licensing Examination 2008). Each step of the USMLE emphasizes a graded understanding of basic and clinical science principles, leading to the eventual application of learned principles to solving treatment and patient management scenarios under assumed clinical independence.

The USMLE is a multiple-choice (was paper-pencil and is now computer based) test, and scores across all steps range between 140 and 280. The passing score is 182 for Steps 1 and 2, and 184 for Step 3 (United States Medical Licensing Examination 2008). Mean scores for all of the stages of the USMLE are between 200 and 220 with a standard

deviation of 20 (United States Medical Licensing Examination 2008). The USMLE is an important outcome variable in establishing the “fit” of an individual into the practice of medicine as it attempts to establish, through measuring various basic and higher-order thought processes, clinical readiness related to knowledge.

Historically, two principal factors predicting successful performance on USMLE Steps 1 and 2 have been UGPA and MCAT subtest scores (particularly science subtests). Step 3 of the USMLE which is heavily weighted on assessing the application of clinical reasoning skills is not as strongly predicted by MCAT and UGPA (United States Medical Licensing Examination 2008). However, the MCAT subtest observed to most predict performance on USMLE Step 3 has been the verbal reasoning subscale. One of the explanations behind this result involves the idea that the verbal reasoning subtest, which assesses the organization and application of verbal skills and Step 3 which assesses the application of clinical reasoning skills (which include the ability to verbalize information to patients concerning issues around diagnosis and treatment)—have overlapping variability. In other words, both tests contain content and require skills that are parallel. Similarly, the association between MCAT science sections and UGPA and Steps 1 and 2 of the USMLE, overlap in the skills they are testing, namely assessing the fund of basic science principles and concepts.

Other variables correlated with performance in medical school and beyond undergraduate grade point average (UGPA)

One of the primary objectives of admissions committees is to quantify all pre-admission information in order to calibrate the “quality” and potential of the candidate. Other than the MCAT, other institutional selectivity variables have been used to predict candidate selection and performance in medical school (Dawson-Sauders et al. 1986). The most conventional pre-admission variables known to correlate with the MCAT and predict future medical school success are, undergraduate science GPA and undergraduate total GPA. GPA’s are typically balanced with MCAT as they serve as an index of previous achievement capacity and a measure of scholastic aptitude. Many studies have documented not only the relationship between UGPA [particularly that science GPA shares with MCAT scores] but also their significant unique contribution in accounting for variability in outcomes measures such as, medical school GPA (MGPA) and Steps 1, 2, and 3 of the USMLE. Moreover, studies have shown that when UGPA is combined with subscales of the MCAT (particularly biological sciences and verbal reasoning) in predicting medical school success and licensure performance, the validity coefficient is further magnified. In one study, results from a regression analyses revealed that total UGPA accounted for 13% of clerkship performance, MCAT scores accounted for 21% (incremental validity of the MCAT), however combined both variables in the prediction algorithm accounted for 28% of the variability in clerkship performance (Huff et al. 1999).

When considering single predictors science subscales of the MCAT are better predictors of medical school and licensure performance than UGPA alone (Crowder 1959; Feil et al. 1998). This finding may be a result of the fact that content within science sections of the MCAT more closely mirror the content covered in the first couple of years in medical school compared to the more variable content covered in an undergraduate degree. Nonetheless, UGPA has been found to be an important variable of later success and was included as an observed variable in the present study (Veloski et al. 2000).

Age

Age at matriculation (i.e., entering into medical school) has also been studied to be a significant predictor in differentiating those who do well from those who do not, and also in predicting performance in pre-clinical and clinical years of medical school. Studies have shown that older (between 25–29 and 30 and above) medical students perform less well in medical school and on licensure exams than their younger (between 22 and 24) counterparts. One group of researchers from McGill university examined differences between younger and older medical students to investigate potential age disparities (Feil et al. 1998). The results showed that older students had lower GPA and MCAT scores and were interviewed 44% less frequently than younger students (Feil et al. 1998). As well, during the first 2 years of medical school (basic science years) test scores were lower for older students than for younger students (Feil et al. 1998). However, this difference dissipated as both older and younger medical students progressed through their educational program into their clinical years. In terms of clinical performance, results revealed that there were no significant differences between older and younger medical students (Feil et al. 1998). Similarly, these results have been replicated in other studies. Although age differences appear to dilute over time, it continues to be an important variable in predicting success in medical school, and is a factor that remains significant in admission into medical school.

Summary

A number of studies have shown that the MCAT has adequate internal consistency reliability (a necessary but not sufficient condition for validity of test scores). Moreover, these studies have provided results from across years and medical institutions in the United States and Canada, supporting the predictive validity (a necessary and sufficient condition for a screening and selection instrument) of the MCAT. Variability in performance on the MCAT accounts for some individual and group differences in performance on pre-clinical and clinical criteria (i.e., USMLE Steps 1, 2, and 3) of success. Such results have been reproduced in many studies, further documenting the predictive power of MCAT, and have also reported on the strength of associations between MCAT subscales and other related cognitive criteria such as undergraduate success (i.e., UGPA) (Tekian et al. 1996). In addition to predictive validity, another important aspect of assessing the validity of a measurement is its correlation with other measurements taken more-or-less concurrently (i.e., concurrent criterion-related validity). As mentioned previously in this paper, the MCAT has been shown to have significant correlations with other cognitive-based assessments such as pre-medical undergraduate GPA, and has been sampled across several medical schools and diverse testee groups from the United States and Canada (Donnon et al. 2007). These findings not only support the concept of concurrent related validity but are further strengthened by the stability of the measurement instrument across time, institution and multiple testee groups.

There is a common theme throughout undergraduate and post-graduate medical education that includes important aspects of a construct—aptitude for medicine—assessed by the MCAT. It is evident from the foregoing discussion that the MCAT is assumed to assess the construct of *aptitude for medicine* and that UGPA is a measure of *general achievement* prior to, during, and subsequent to medical school. Moreover, the combination of these constructs provides the best prediction of *competence in medicine* during medical school in the preclinical years, clerkship years, and beyond into residency than any of the variables alone. Little research has been published testing an entire model of aptitude-achievement-

competence identifying these as latent variables, however. Most research heretofore has employed simple correlation or regression methods.

The primary purpose of the present study, therefore, was to further develop and test a latent variable path model of *general achievement*, *aptitude for medicine* and *competence in medicine*. This latent variable path model, provides a parsimonious model which accounts for maximum variability in indicators of success (e.g., preclinical academic performance, performance in stages on licensure examinations—USMLE Steps 1, 2, and 3) in medical school and beyond. With such a model we not only sought to clarify the predictive validity of the MCAT on future criteria of success, but also to explicate a full measurement and structural model of achievement, aptitude and competence.

Methods

Data sources and sample

The data were acquired from the MCAT offices at the Association of American Medical Colleges (AAMC) and the National Board of Medical Examiners and consisted of 839,710 participants from 1991 to 2000, across 115 medical schools. Data were available on demographic and school variables, UGPA (Undergraduate GPA's before medical school—UGPA's were presented both for each undergraduate year), MCAT subtest scores (biological sciences, physical sciences, verbal reasoning, and writing sample) and, for those who had been admitted to medical school ($n = 250,792$), Steps 1, 2, and 3 of the USMLE. This study received approval from the Conjoint Health Research Ethics Board of the University of Calgary.

Descriptive data

There were more males than females (58.3% and 41.7%, respectively) in the sample, many were between 22 and 24-years-old (58%) at matriculation, the majority were not under-represented minorities (91%) and more than two-thirds scored between 8 and 12 on the MCAT subtests. Males performed significantly better ($p < .05$) than females on the biological sciences, physical sciences and verbal reasoning subtests of the MCAT but not on the writing sample subtest. Similarly, males outperformed females on Step 1 of the USMLE but not on Steps 2 and 3.

Age was inversely related to undergraduate GPA (both science and non-science courses) but not to MCAT subtests scores or USMLE Steps 1–3 scores.

Latent variable path analysis (LVPA)

We employed latent variable path analysis to define and describe the relationships and directional pathways between and among the observed/measured variables (MCAT subscales, UGPA, age, and USMLE) that measured the underlying latent constructs (general achievement, aptitude in medicine and competence in medicine).

Overview of latent variable path modeling

Of the many multivariate statistical methods (e.g., multiple regression, factor analysis), latent variable path analysis (LVPA), a subset of structural equation modeling (SEM), has

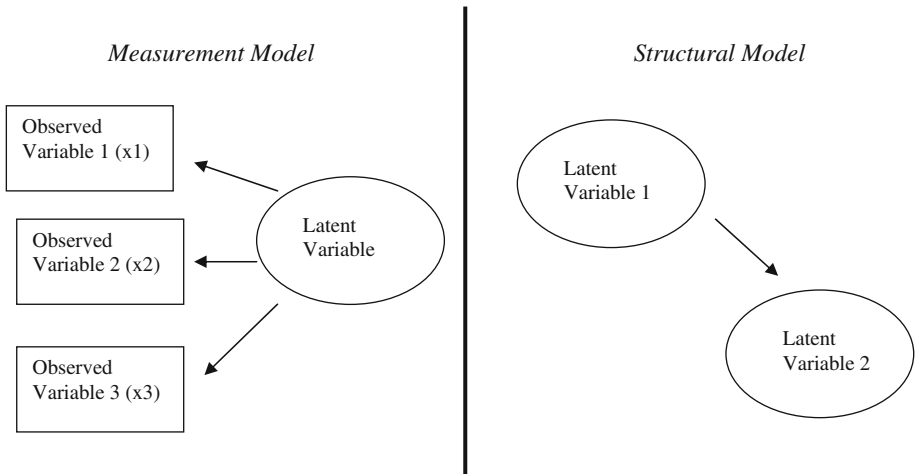


Fig. 1 Illustration of the measurement and structural models subsumed under SEM

become an important research tool for making explicit the theoretical and statistical organization among a set of variables. SEM is a super ordinate term that contains many statistical procedures such as LVPA. It was used in the present study as a general model for explaining its role and function as a theory strong approach to accounting for multivariate variability. SEM builds on some rudimentary statistical analyses such as analysis of variance (ANOVA) and multiple regression, but moves a step beyond to include computational procedures that are able to deduce direct and indirect causal links amongst latent and observed variables (Jöreskog 1982). In this way SEM is a unified system for patterning the structural relationships among variables.

Theory and importance of SEM

SEM is decomposed of two fundamental parts which include: (1) the measurement model and (2) the structural model (refer to Fig. 1). In the measurement model latent variables or hypothetical variables are measured and identified through observed variables (also referred to as indicators of the latent variables). In the present study the latent variables were named *General Achievement* measured by UGPA, *Aptitude for Medicine* measured by subscales of the MCAT, and *Competence in Medicine* measured by Steps 1, 2, and 3 of the USMLE and the MCAT verbal reasoning subscale. The second part of the model involves the structural relationships, which are denoted by the path links and coefficients between and among the latent variables (Jöreskog 1982).

The measurement model involves the assessment of variables assumed to have manifested from an underlying structure. These variables are referred to as observed or measured variables. As it is essential to account for as much variability in the latent variable as possible, many researchers have stressed that using multiple indicators (observed variables) increases the capacity to account for more variance and decreases the capacity for creating biased estimates relating to the structural parameters surrounding the latent construct (Jöreskog 1982). One of the primary objectives of LVPA is to fit the predicted model [that has been theorized] to the observed data [measured by the observed variables]. The process of achieving the above outcome is determined by ensuring the

inclusion of adequate model characteristics in order to minimize misspecification of the model. Therefore, in order to develop and test a robust LVPA, the following are the criteria to be considered: (Raykov 2000; Myung 2003; Bollen 1989; Bentler 1982).

1. A well developed theory postulating a testable model is established;
2. Model specification—hypothesis about all variables to be included in the model are operationalized and clearly specified;
3. Selecting appropriate and multiple indicators or observed measures;
4. Ensuring psychometric features of measurement scales used to test observed variables are robust (reliabilities and validities);
5. Large sample size;
6. Use of longitudinal data to ensure stability of the model;
7. Proper use of estimation procedures—model parameters of the predicted model are compared to those of the observed model (e.g., maximum likelihood estimation);
8. Proper use and interpretation of Fit Indices—used to verify the degree to which a model fits the data (e.g., χ^2 , comparative fit index (CFI), root mean square error of approximation (RMSEA)). A rule of thumb for determining whether or not there is good fit is a fit $>.90$;
9. Respecification of the model [if needed].

Data analyses

Based on all of the previous results and theoretical considerations, we constructed an LVPA model of *general achievement*, *aptitude for medicine* and *competence in medicine* as depicted in Fig. 2. This LVPA with these three latent variables was tested on a large sample ($\sim 10\%$; $n = 20,714$).

We employed the recommended rules for fit indices in latent variable path models including Bentler's comparative fit index (CFI) (Bentler 1982) with maximum likelihood estimation, the standardized root mean squared residual (SRMR), and the root mean squared error of approximation (RMSEA).

A model was tested and ten variables (UGPA Year 1, UGPA Year 2, UGPA Year 3, UGPA Year 4, MCAT Bio Sc, MCAT Phys Sc, MCAT Verbal, USMLE Step 1, USMLE Step 2, USMLE Step 3) were included as identifying the three latent variables (general achievement, aptitude for medicine, competence in medicine). The MCAT written subtest was not included in this analysis since our initial regression analyses indicated that it correlated $R = .0$ with all the variables. This same finding has been documented in other studies (Donnon et al. 2007; Swanson et al. 1996). Age was also included in the model because it was included in the multiple regression analyses and contributed significant variance, as well as retained notable loadings in the factor analysis. These ten measured variables and one demographic variable (age), were intercorrelated and converted to a variance-covariance matrix fitting the model using maximum likelihood (ML) estimation.

The overall fit of the model to the data was good producing a CFI = .932, a $\chi^2 = 3209.70$, $p < .001$, and a RMSEA = .092. The minimizing of the Q function proceeded smoothly requiring five iterations. A CFI = .932 indicates that 93.2% of the variance and covariance in the data is accounted for by the proposed model.

The path coefficients and other parameters are summarized in Fig. 2. Path coefficients can be interpreted as standardized partial coefficients (β) (analogous to regression coefficients in a multiple regression). All three latent variables are clearly identified. *General*

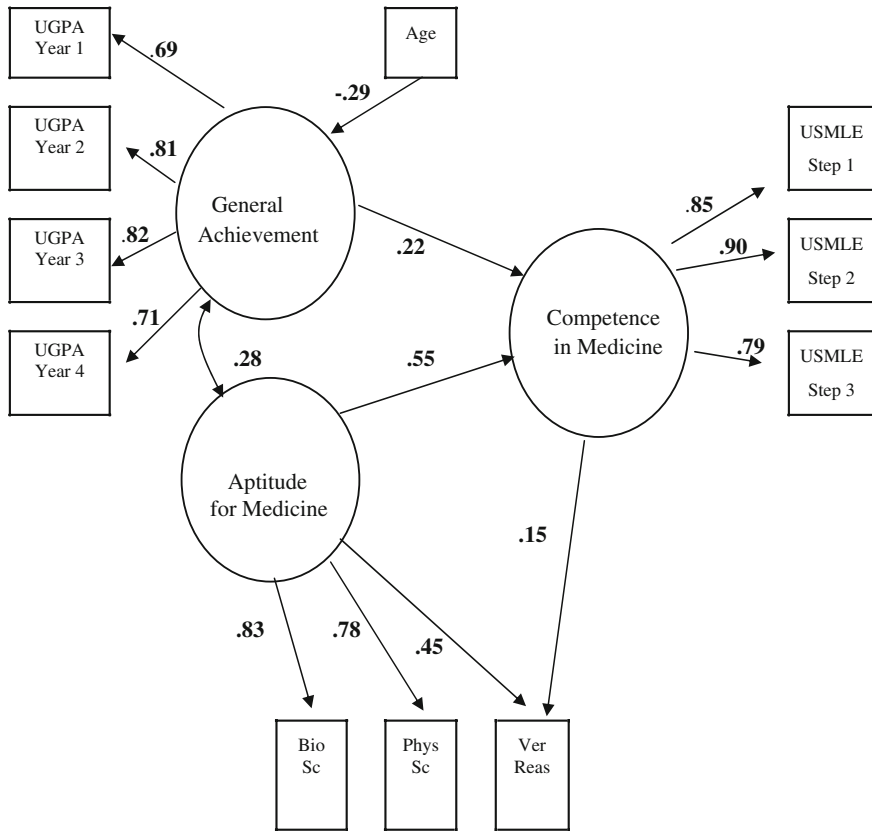


Fig. 2 Latent variable path model of undergraduate GPA, MCAT, USMLE steps, and age employing maximum likelihood estimation ($n = 20,714$); Model fit statistics: $\chi^2(41) = 3209.70$, $p < .001$; comparative fit index = .932; root mean squared error of approximation = .092

achievement (path coefficients ranging from .69 to .82) is identified by four variables (UGPA Year 1, UGPA Year 2, UGPA Year 3, UGPA Year 4) that are all theoretically relevant to an achievement construct. The path coefficients on *Aptitude for medicine* range from .45 to .83 and are measured by three subtests of the MCAT (Bio Sc, Phys Sc, Verbal). Three of the measured variables (USMLE Step 1, USMLE Step 2, USMLE Step 3) also serve to identify *Competence in medicine*. There is a direct path coefficient (.55, $p < .001$) from aptitude in medicine to competence in medicine that confirms the “causal” direction. The correlations between general achievement ($r = .22$, $p < .001$) and competence in medicine and aptitude for medicine ($r = .28$, $p < .001$) all confirm the expected relationship.

The *General achievement* factor has a significant path coefficient ($-.29$, $p < .01$) from age (this path accounts for 8.4% of the variance) indicating that age is related to undergraduate achievement (inversely—younger students perform better than older ones). The split loading for the verbal reasoning subscale of the MCAT, that has a path coefficient on *Aptitude for medicine* (.45) and a path coefficient on *Competence in medicine* (.15) suggests that skills tested on the MCAT verbal reasoning subscales and portions of Step 3 of the USMLE have overlapping variability.

All of the latent variables are related as predicted, and the crucial path coefficient from aptitude in medicine to competence in medicine is significant (path coefficient = .55; $p < .001$). The intercorrelations among the other latent variables are all significant. Overall, these results support the overall model and the particulars of the predicted relationships as well.

Discussion

The main finding of the present study is that a latent variable path model of general achievement, aptitude for medicine and competence in medicine fit a sample of the data well, with a CFI = .932. This result provides confirmatory evidence for construct validity and defined three latent variables that can be identified as separate but related entities as theoretically expected (correlated and significant path coefficients).

It is clear that many of the variables employed heretofore and used in the present study share some degree of relationship (i.e., collinearity) with each other. Since the selection process for medical school is focused on evaluating cognitive-based or achievement related aptitude, the underlying structure of these variables appears to be similar in that they sample cognitive ability. MCAT scores and UGPA are better predictors of the first 2 years of medical school than of later years, probably because much of the content is declarative biomedical knowledge in all of the assessed variables. As students move through medical school into residency, these associations steadily decrease probably because different skills and abilities (e.g., clinical reasoning judgment), are employed in order to effectively manage learning at these higher levels of education (p. 31).

A limitation of the present study was that the data consisted primarily of cognitive variables (i.e., achievement, aptitude, competence). It would be beneficial to include other important noncognitive measures such as volunteer experience, personality characteristics, social-emotional intelligence and motivational factors, and so forth in order to produce a more comprehensive model of undergraduate and post-graduate medical education success and competence. Future research might well include these variables so that a more comprehensive LVPA model than the present one may be tested.

Notwithstanding this limitation, there were several strengths of the present study. First, a very large database was employed ($n > 800,000$). Second, the data were longitudinal over a 10-year period thus adding to the stability of the findings. Third, complex and advanced statistical analyses (i.e., LVPA) were employed to test several research questions and comprehensive models. Fourth, the data were based on many American and Canadian medical schools (115) and the results cannot, therefore, be attributable to one or several unique institutions. Fifth, data were available on pre-medical school achievement (i.e., UGPA), aptitude (i.e., MCAT scores), and pre-clinical and clinical performance in medicine (i.e., USMLE Steps 1, 2, and 3). This allowed the construction of the achievement-aptitude-competence model. Given these strengths we can feel confident about the stability and generalizability of the present findings.

Summary and conclusion

The present results not only provided evidence for the predictive validity of the MCAT (with an incremental increase with undergraduate GPA) but also for the overall LVPA model positing structural relationships between and among *general achievement*, *aptitude*

for medicine, and competence in medicine. The principal question addressed by most research continues to be: “What characteristics of candidates predict the success or performance of a medical student and resident?” We can say that the present study answered part of this question—that is cognitive-achievement related variables are an important aspect of success throughout both undergraduate and post-graduate medical education.

Notwithstanding the MCAT’s predictive validity, there have been many calls recently (Parlow and Rothman 1974; Cohen 2001) to explore factors other than the MCAT—particularly non-cognitive attributes associated with effective physicians—as potential criteria for selection into medical school. These may include key personal characteristics such as altruism, empathy, integrity and compassion (Cohen 2001). Important medical professional organizations such as the Accreditation Council for Graduate Medical Education (ACGME), the American Board of Medical Specialties (ABMS), and the Royal College of Physicians and Surgeons of Canada have emphasized the multiplicity of physician roles such as medical expert, collaborator, manager, health advocate, scholar, professional and communicator (The Medical School Objectives Writing Group 1999; Societal Needs Working Group 1996; Accreditation Council for Graduate Medical Education 2001). A challenge to the medical profession, then, is to develop screening and selection methods and devices that combined with instruments like the MCAT, can closer approximate the potential of those applying to medical school and moving closer to independent medical practice.

Acknowledgements We would like to thank the Association of American Medical Colleges and the National Board of Medical Examiners for providing access to the data that allowed us to conduct this study. In particular we would like to thank the following people for their support and help: Ellen Julian, Pat Etienne and Robert Galbraith.

Funding The first author received the 2004 MCAT Graduate Student Research Program (GSRP) award to support this research under the supervision of the second author.

Contributors The first and second authors contributed to the development and execution of this research. All three authors contributed to the data analysis, and the writing and editing of the manuscript.

References

- Accreditation Council for Graduate Medical Education. (2001). *Graduate medical education directory 2001–2002*. Chicago, IL: AMA.
- Albanese, M. A., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing personal qualities in medical school admissions. *Academic Medicine*, 78, 313–321. doi:10.1097/00001888-200303000-00016.
- Association of American Medical Colleges. (1991). *Use of MCAT data in admissions: A guide for medical school admissions officers and faculty*. Washington, DC: AAMC.
- Basco, W., Way, D., Gilbert, G., & Hudson, A. (2002). Undergraduate institutional MCAT scores as predictors of USMLE step 1 performance. *Academic Medicine*, 77, s13–s15.
- Bentler, P. M. (1982). Linear systems with multiple levels and types of latent variable. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation—causality—structure—prediction* (pp. 1–18; 102–130). New York: North-Holland Publishers.
- Bollen, K. A. (1989). *Structural equations with latent variables* (pp. 1–80; 226–236). New York: Wiley.
- Cohen, J. (2001). Facing the future. President’s Address. In *112th Annual Meeting of the Association of American Medical Colleges*. Washington, DC, November 4, 2001.
- Crowder, D. G. (1959). Prediction of first-year grades in a medical college. *Educational and Psychological Measurement*, 19, 637–639. doi:10.1177/001316445901900418.
- Dawson-Saunders, B., Paiva, R. E. A., & Doolen, D. R. (1986). Using ACT scores and grade point averages to predict students’ MCAT scores. *Journal of Medical Education*, 61, 681–683.

- Donnon, T., Oddone, E., & Violato, C. (2007). The predictive validity of the MCAT on medical school performance and medical board licensing examinations: A meta-analysis of the published research. *Academic Medicine* (in press).
- Donnon, T., & Violato, C. (2006). Medical students' clinical reasoning skills as a function of basic science achievement and clinical competency measures. *Academic Medicine*, *81*, S12–S15. doi:[10.1097/01.ACM.0000236543.88782.b6](https://doi.org/10.1097/01.ACM.0000236543.88782.b6).
- Feil, D., Kristian, M., & Mitchell, N. (1998). Older medical students' performances at McGill University. *Academic Medicine*, *73*, 98–100. doi:[10.1097/00001888-199801000-00020](https://doi.org/10.1097/00001888-199801000-00020).
- Gilbert, G. E., Basco, W. T., Jr., Blue, A. V., & O'Sullivan, P. S. (2002). Predictive validity of the Medical College Admissions Test writing sample for the United States Medical Licensing Examination Step 1 and 2. *Advances in Health Science Education*, *7*, 191–200. doi:[10.1023/A:1021110132214](https://doi.org/10.1023/A:1021110132214).
- Hojat, M., Erdmann, J. B., Veloski, J. J., et al. (2000). A validity study of the writing sample section of the Medical College Admission Test. *Academic Medicine*, *75*, S25–S27. doi:[10.1097/00001888-200010001-00008](https://doi.org/10.1097/00001888-200010001-00008).
- Huff, K. L., Koenig, J. A., Treptau, M. M., & Sireci, S. G. (1999). Validity of MCAT scores for predicting clerkship performance of medical students grouped by sex and ethnicity. *Academic Medicine*, *74*, S41–S44.
- Jones, R. F., & Thomae-Forgues, M. (1984). Validity of the MCAT in predicting performance in the first two years of medical school. *Journal of Medical Education*, *59*, 455–464.
- Jöreskog, K. G. (1982). The LISREL approach to causal model-building in the social sciences. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation—causality—structure—prediction* (pp. 1–18; 81–100). New York: North-Holland Publishers.
- Koenig, J., Sireci, S., & Wiley, A. (1998). Evaluating the predictive validity of MCAT scores across diverse applicant groups. *Academic Medicine*, *73*, 1095–1106. doi:[10.1097/00001888-199810000-00021](https://doi.org/10.1097/00001888-199810000-00021).
- Mitchell, K., Haynes, R., & Koenig, J. A. (1994). Assessing the validity of the updated Medical College Admission Test. *Academic Medicine*, *69*, 394–401.
- Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*, 90–100. doi:[10.1016/S0022-2496\(02\)00028-7](https://doi.org/10.1016/S0022-2496(02)00028-7).
- Parlow, J., & Rothman, A. I. (1974). Personality traits of first year medical students: trends over a six year period. *British Journal of Medical Education*, *54*, 759–765.
- Raykov, T. (2000). On sensitivity if structural equation modeling to latent relation misspecifications. *SEM*, *7*, 596–607.
- Societal Needs Working Group. (1996). CanMEDS 2000 project. Skills for the new millennium. *Annals of the Royal College of Physicians and Surgeons of Canada*, *29*, 206–216.
- Swanson, D. B., Case, S. M., Koenig, J. A., & Killian, C. D. (1996). Preliminary study of the accuracies of the old and new Medical College Admission Tests for predicting performance on USMLE Step 1. *Academic Medicine*, *71*, S25–S30.
- Tekian, T., Mrtek, R., Syftestad, P., Foley, R., & Sandlow, L. (1996). Baseline longitudinal data of undergraduate medical students at risk. *Academic Medicine*, *71*, s86–s87. doi:[10.1097/00001888-199610000-00053](https://doi.org/10.1097/00001888-199610000-00053).
- The Medical School Objectives Writing Group. (1999). Learning objectives for medical student education—guidelines for medical schools: Report I of the medical school objectives project. *Academic Medicine*, *74*, 13–18.
- United States Medical Licensing Examination. (2008). *Scoring and score reporting*. Philadelphia, PA: NBME.
- Veloski, J. J., Callahan, C. A., Xu, G., Hojat, M., & Nash, D. B. (2000). Prediction of students' performances on licensing examinations using age, race, sex, undergraduate GPA, and MCAT scores. *Academic Medicine*, *75*, S28–S30. doi:[10.1097/00001888-200010001-00009](https://doi.org/10.1097/00001888-200010001-00009).
- Violato, C., & Donnon, T. (2005). Does the Medical College Admission Test predict clinical reasoning skills? A longitudinal study employing the Medical Council of Canada Clinical Reasoning Examination. *Academic Medicine*, *80*, S76–S78. doi:[10.1097/00001888-200510001-00007](https://doi.org/10.1097/00001888-200510001-00007).