Assessing Personal Qualities in Medical School Admissions

Mark A. Albanese, PhD, Mikel H. Snow, PhD, Susan E. Skochelak, MD, MPH, Kathryn N. Huggett, MA, and Philip M. Farrell, MD, PhD

ABSTRACT

The authors analyze the challenges to using academic measures (MCAT scores and GPAs) as thresholds for admissions and, for applicants exceeding the threshold, using personal qualities for admission decisions; review the literature on using the medical school interview and other admission data to assess personal qualities of applicants; identify challenges of developing better methods of assessing personal qualities; and propose a unified system for assessment.

The authors discuss three challenges to using the threshold approach: institutional self-interest, inertia, and philosophical and historical factors. Institutional self-interest arises from the potential for admitting students with lower academic credentials, which could negatively influence indicators used to rank medical schools. Inertia can make introducing a new system complex. Philosophical and historical factors are those that tend to value maximizing academic measures.

The literature identifies up to 87 different personal qualities relevant to the practice of medicine, and selecting the most salient of these that can be practically measured is a challenging task. The challenges to developing better personal quality measures include selecting and operationally defining the most important qualities, measuring the qualities in a cost-effective manner, and overcoming "cunning" adversaries who, with the incentive and resourcefulness, can potentially invalidate such measures.

The authors discuss potential methods of measuring personal qualities and propose a unified system of assessment that would pool resources from certification and recertification efforts to develop competencies across the continuum with a dynamic, integrated approach to assessment.

Acad. Med. 2003;78:313–321.

Making decisions about whom to admit to medical school has come under increasing scrutiny. Jordan Cohen, MD, in his address to the 112th annual meeting of the Association of American Medical Colleges, decried the use of undergrad-

uate grade-point averages (GPAs) and Medical College Admission Test (MCAT) scores as the primary means of selecting medical students. Instead of using these indicators, he suggested using "MCAT scores and GPAs only as threshold measures," or "beginning the screening with an assessment of personal characteristics and leave the GPAs and MCAT scores 'til later." He argued that using this approach "admission committees might well find many instances in which truly compelling personal characteristics would trump one or two isolated blemishes in the academic record."[1]

With Dr. Cohen's challenge to look first to personal qualities in the admission process, the need to effectively measure personal qualities has assumed greater importance. In this article, we review the literature on the medical school interview as a mechanism for assessing personal qualities, discuss the challenges in using the interview and other

Dr. Albanese is professor of population health sciences and director, Office of Medical Education Research and Development; Dr. Snow is professor of anatomy and associate dean for Students and Admissions; Dr. Skochelak is professor of family medicine and senior associate dean for Academic Affairs; Ms. Huggett was director, Medical Scholars Program, at the time this study was conducted; and Dr. Farrell is vice chancellor for medical affairs, dean, and Alfred Dorrance Daniels Professor on Diseases of Children; all are at the University of Wisconsin–Madison Medical School, Madison, Wisconsin.

Correspondence and requests for reprints should be addressed to Dr. Albanese, Office of Medical Education Research and Development, University of Wisconsin Medical School, 1300 University Avenue—Room 2070 MSC, Madison, WI 53706-1532.

approaches to assessing personal qualities, and then suggest approaches that might be taken to address this important issue.

BACKGROUND

Historically, the interview has been one of the primary methods of assessing personal qualities. Interviews for admission to medical school are conducted by all but a few U.S. medical schools.[2,3] Edwards et al.[3] cite four purposes for the admission interview: information gathering, decision making, verification, and recruitment. They argue that the most important purpose of the interview is to gather nonacademic information about candidates that would be difficult or impossible to obtain by other means. The method they advocate for obtaining this information is a Success Analysis of Medical Students (SAMS), which includes selecting interview content based upon a job analysis (the critical-incidents technique is advocated for this purpose), standardizing the questions asked of all applicants, providing interviewers with sample answers to questions to help them give consistent ratings, and conducting each interview with a board or panel of interviewers. Interviews have been classified as being "structured" (like the SAMS model), semi-structured (having some but not all elements of a structured model), and "unstructured."

Substantial evidence exists that admission committees place great emphasis on the information gleaned from interviews. Puryear and Lewis[2] reported that 61% of 107 medical schools responding to their survey stated that the admission interview data were the most important variables used in selection. Empirical data support this result as well. Nowacek et al.[4] found that, after interviewing candidates, admission committee members changed their ratings of the candidates, with mean values for various assessed qualities changing by as much as .47 of a standard deviation (effect size or ES). Patrick et al.[5] reported the impact of introducing interview data obtained using the SAMS model on admission decisions. After adding the interview ratings to information from the written application, the percentage of variance in acceptance decisions accounted for by the regression model increased from 21% to 37%. Data obtained from the admission interview clearly can have a significant effect on admission decisions, but what are the non-academic qualities being assessed in the interview and in what ways are they being assessed?

Meridith et al.[6] reported rating an applicant's maturity, individual achievement, motivation/interest in medicine, ability, and interpersonal skills. Nowacek et al.[4] evaluated communication and interpersonal skills, commitment to serve others, familiarity with issues in medicine, leadership ability, motivation for medicine, and overall impression. Murden et al.[7] assessed applicants' levels of maturity,

nonacademic achievement, motivation, and rapport. Powis et al.[8] assessed perseverance, tolerance of ambiguity, supportive and encouraging behavior, motivation to become a doctor, self-confidence, compatibility with the school's study styles, and an overall judgment. Taylor[9] reported drawing traits assessed in a written form from 87 positive qualities of successful physicians. Collins et al.[10] assessed communication, maturity, caring qualities/friendliness, awareness of community, political, social and medical issues, certainty of career choice, involvement in school activities, and involvement in community activities. Shaw et al.[11] assessed 20 "non-cognitive, non-teachable traits," such as being honest, energetic, confidence-inspiring, and conscientious. These authors are not alone in their beliefs that certain noncognitive traits are non-teachable. Bullimore[12] argues that personality is set by age 18, making assessment of noncognitive variables in the medical school admission interview critical. The concept of personality and non-teachable traits implies traits that are stable across time and situations. Is honesty really non-teachable? Is an energetic person always energetic? Even if one accepts that there are some non-teachable traits, might they not be coachable for display in a one- to two-hour interview?

Admission to medical school is a high-stakes proposition for almost all applicants. Many applicants spend significant sums of money for test preparation services for the MCAT. Might a standardized interview purporting to assess non-teachable skills find itself susceptible to coaching from such a service? Further, is an interview the only way or even the best way to assess these noncognitive traits? Taylor[9] argues that such traits can be assessed by having candidates distribute evaluation forms to individuals of their choice as part of the application process. He further reports that the students selected at the University of Iowa College of Medicine using such an approach did not differ from those selected when an traditional interview was conducted. In the application process at the University of Wisconsin Medical School, the personal statement has served as a key indicator of noncognitive traits. Our literature review, however, found no study that examined to what extent admission committees' assessments of the personal statement yielded different assessments of applicants' noncognitive qualities than an interview.

Ultimately, the decision whether an interview is worth the time and expense must be based on whether the interview yields something that cannot be obtained by other means, and, in particular, something that cannot be obtained from a review of written application materials. Evidence for the value of the interview has been sought in studies assessing the reliability and validity of the interview. The results have been equivocal. Studies of the reliability of interviewers have produced quite variable estimates. Meridith et al.[6] found inter-rater correlations ranging from .55 to .91 for five

qualities assessed in a sample of 14 applicants, each evaluated by two raters. Powis et al.[8] report kappa (chance corrected inter-rater agreement) statistics ranging from .23 to .63 for seven qualities independently assessed by two raters. Edwards et al.[3] report results from several meta-analyses showing inter-rater reliabilities ranging from .52 to .96, with a median of .83. Reliabilities for studies using structured interviews ranged between .82 and .84, and for those using unstructured interviews, reliabilities ranged from .61 to .75. Nowacek et al.[4] reported inter-rater reliabilities for overall impressions of applicants that were .57 before the interview and .55 after the interview. Richards et al.[13] reported an inter-rater reliability of .67 for panels of 13 interviewers. Van Susteren et al.[14] reported an inter-rater kappa reliability of .79 for interviewers providing ratings scored within one point of each other on a five-point scale. Inter-rater reliabilities appear to be quite variable, but generally were higher ($>.8$) for structured interviews.

The validity of interviews has also proven equivocal. Litton-Hawes et al.,[15] analyzing 15 interviews using stimulated recall procedures from videotapes, found interviewers made inefficient use of time and focused on written materials to the detriment of exploring what they were intended to do. They advocated improved training of interviewers. Smith et al.[16] compared first-year medical students' grades for two classes that had been interviewed with those of two classes that had not been interviewed ($n = 44$ and 79, respectively). Results showed no difference in grades. Perhaps the study producing the most compelling results in support of interview data for admissions comes from Powis et al.[8] In a case–control study designed to retrospectively analyze differences between students who left medical school due to failure or withdrawal over a nine-year period and students who received honors, 56 paired cases (who left medical school) and controls (who completed medical school and who were matched according to gender, age, and entry cohort—all had excelled in their academic performances) were analyzed. Those who left had uniformly been rated more poorly in the interview, with effect sizes of -4.17 for supportive and encouraging behavior, -3.46 for assessments of self-confidence and motivation to become a doctor, -3.11 for the overall rating, -2.76 for compatibility with study style of school, -1.98 for perseverance, and $-.97$ for tolerance of ambiguity. For differences between 58 pairs of students who graduated with honors and matched controls, honors graduates were rated more positively for perseverance, ES = 2.98; self confidence, ES = 2.59; overall rating, ES = 2.17; tolerance of ambiguity, ES = 1.04; and supportive and encouraging behavior, ES = .86. For the remaining qualities, honors recipients received more positive evaluations (with ES < .40). The ES values in this study are meaningful and strongly suggest that interview ratings can discriminate between students who fail to complete medical school and those who complete medical school, as well as between those who graduate with honors and those who do not.

Meridith et al.[6] also provided compelling evidence to support conducting admission interviews. They correlated data collected from the admission interview, as well as the MCAT score and undergraduate GPA, with National Board of Medical Examiners (NBME) Part II scores and subjective clinical assessments in pediatrics and internal medicine clerkships for third-year medical students. Admission interview data did not significantly correlate with NBME Part II scores, but did correlate with the subjective clinical assessments, accounting for over twice the variance as the next most potent predictor (interview = 10.4%, MCAT Science—Quantitative = 5.0%). Similar correlations of interview assessments with clinical assessments but not academic performances have been found in several studies involving non-medical health sciences programs.[17–19] Thus, evidence exists that the interview provides information for admission related to students' performances in the clinical portions of medical education.

It is extraordinary that, among all the reasons given for conducting an admission interview, we found no mention of its use to demonstrate that a school values the personal interaction between human beings, that the admission process is not just a mechanical analysis of paper credentials and accomplishments but a judgment of one's qualities as a human being and a future colleague, particularly because the physician–patient relationship can be so intensely personal. The interview can also be a means of demonstrating compassion for applicants whose records may have temporary performance deficits that may be related to deaths in the family, illness, or other problems. The interview is a chance for an institution to place a human touch on what is a highly stressful, high-stakes decision process for all involved.

To summarize, the AAMC has called for a greater emphasis on compelling personal characteristics in the selection of medical students. These compelling personal characteristics are the focal point of admission interviews that are conducted by all but a few medical schools in the United States. Evidence exists that admission committees give substantial weight to interview data in the selection of applicants, but what constitutes a compelling personal characteristic varies among institutions, with as many as 87 different qualities being considered for assessment. Even though the evidence for the validity of the interview has been equivocal, there is evidence that interview ratings are predictive of subjective clinical assessments, and low interview assessments are predictive of failure or withdrawal from medical school. The reliability of the interview can be improved using structured approaches.[3] Some have argued there are more cost-effective methods than the interview for assessing compelling personal characteristics.

CONSIDERING ALTERNATIVES

Because interviews are expensive for both the applicant and the institution, it is this issue we address in the remainder of this article. Our goal is to analyze the challenges in measuring compelling personal characteristics and then offer some practical and some perhaps less practical alternatives.

Challenges to Reducing Reliance on MCATs and Undergraduate GPAs

The major challenges facing any school adopting Dr. Cohen's recommendation to use a minimum GPA and MCAT score as a threshold *and* measures of compelling personal characteristics for admission are self-interest, inertia, and philosophical and historical factors.

Self-interest. Perhaps the biggest challenge a medical school faces in reducing reliance on academic credentials in admissions is the impact such a reduction may have on the perceptions of others. Mean MCAT scores and undergraduate GPAs are used as part of the formula in determining the "best" medical schools by *U.S. News and World Report.* Further complicating the situation, several studies[20,21] have reported that MCAT scores correlate fairly strongly with United States Medical Licensure Examination (USMLE) Step 1 scores (multiple correlations of MCAT scores with NBME Part I, predecessor of Step 1, between .39 and .63, median = .58[20]; multiple correlation = .59[21]). Since a correlation can reflect either a cause-and-effect relationship (unlikely in this case) or the influence of a third variable (say academic or test-taking aptitude), a reduction in MCAT scores may put USMLE Step 1 scores at risk. Even if failure rates do not rise, lowered mean USMLE Step 1 scores can have substantial damaging effects on an institution. In our experience, top applicants commonly ask us for our mean USMLE Step 1 score, ostensibly a factor they are considering in making their medical school selections. Thus, a lower USMLE Step 1 mean score has the potential to damage recruitment efforts. Compounding the problem, some of the most competitive residency programs consider USMLE Step 1 scores in their decisions. Thus, from the standpoint of various outcomes used to assess the quality of medical schools, ignoring academic credentials beyond a low threshold will bump up against self-interest. On the other hand, setting a high threshold may cripple efforts to identify students with the compelling personal characteristics that may be most prized.

Inertia. Admission is a high stakes, big business operation involving a large number of very eclectic individuals. Although change never comes easily, it is especially difficult under these circumstances. Coming to agreement about what constitutes a reasonable threshold will take a substantial and compelling argument. To arrive at threshold values that were acceptable to our admission committee, we analyzed performance data over a 12-year period in which we simulated various thresholds and the resulting impacts on the likelihood of first-time USMLE Step 1 and Step 2 passage and medical school graduation. Thresholds were adopted for which the likelihood of USMLE Step 1 and Step 2 passage and graduation did not improve with higher GPAs and MCAT scores.[22] Other inertia lies with changing the process by which admission occurs. Admission staff have a tough job and take pride in their accomplishments. For staff feeling overworked under the current admission system, doing things differently could seem overwhelming. They may also feel personally threatened by change. Thus, it takes a concerted effort not only to get faculty buy-in but also to ensure that the administrative staff supports the changes.

Philosophical and historical factors. Finally, philosophical and historical factors are likely to be major obstacles. Some faculty believe that we should admit only the "best and the brightest" by academic measures. Reliance on non-academic measures beyond a low threshold would be an anathema from this perspective. The related issue is that a history of encountering problems with students who have low academic credentials can come back to haunt any effort to change. All it takes is one or two such students admitted under the new system encountering major academic problems for faculty to develop resistance to assuming additional risk. The only way to counter these historical and philosophical differences is to collect data on the performances of students in various risk categories (if there are such data available). In this way, the poor outcomes with one or two students can be put in perspective if there have been good outcomes with 30 or 40. If there are appropriate data and the outcomes have not been compellingly positive, one must be prepared to assess the risk and the benefits.

Challenges in Measuring Compelling Personal Characteristics

If a faculty decides that it is willing to take the risk of using a threshold approach for screening applicants and then admitting them on the basis of compelling personal characteristics, it still faces the daunting task of reliably and validly measuring these qualities. Among the challenges are determining: What constitutes a compelling personal characteristic, and which is/are most compelling? What is/are best method(s) of measuring these qualities? To what extent are these qualities influenced by nature, nurture, or maturation? What are the costs of measuring these qualities? What are the ways of overcoming cunning adversaries?

What constitutes a compelling personal characteristic, and which is/are most compelling? What are these compelling personal characteristics that might trump other indicators? Although the literature offers insights into some

qualities that have been assessed by interview and other means, there clearly is room for research into what are the most salient qualities. The 87 positive qualities of successful physicians identified by Price et al.[23] might be a good starting point, but that number of qualities makes it a daunting starting point and, most likely, impractical for measuring. An effort to improve assessment methods for prospective medical students by the AAMC in the early 1970s might be of some help in this regard. Based upon the work of Price et al.,[23] the Non-Cognitive Working Group, under the leadership of Jack Collwell, proposed specific objective measures of seven personal qualities be incorporated into the MCAT: compassion, coping capabilities, decision making, interprofessional relations, realistic self-appraisal, sensitivity in interpersonal relations, and staying power—physical and motivational.[24] Although the recommendations from this working group were never acted upon, the set of personal qualities they identified might contribute usefully to the dialogue about the most salient personal characteristics to assess during the medical school selection process. A lot has transpired in the quarter century since these recommendations. Whether the same personal qualities would be identified today cannot be determined. It might be worth convening a similar working group to update the work or, perhaps, use the nationally directed multi-institutional process employed for the Medical School Objectives Project (MSOP) to update the recommendations. Whatever the recommendations would be, they would still need to be assessed by each institution for local relevance. However, it would be of substantial help to have a nationally defined set as a starting point rather than having each institution develop its own.

What is/are the best method(s) of measuring these qualities? Measuring compelling personal characteristics is challenging for at least two reasons. First, measuring a personal quality requires the difficult step of defining the personal quality in measurable terms. This involves defining the personal characteristic not only in behavioral terms but also in behavioral terms that most reasonable people would recognize as reflecting the personal quality if they were to see it. This is not an easy thing to do. Take altruism as an example. The MSOP delineated seven qualities of altruism that medical students must demonstrate before graduation to the satisfaction of the faculty. The last of these qualities is "The capacity to recognize and accept limitations in one's knowledge and clinical skills, and a commitment to continuously improve one's knowledge and ability."[25] This is not one of those components of altruism that might be in everyone's definition of altruism. Even if it were a generally agreed upon element of altruism, developing an assessment method whereby most people would recognize it if they saw it offers multiple challenges.

The second reason is that one is most interested in stable qualities that have a high probability of occurrence in an almost infinite number of different situations. Take the example of altruism again. Most admission committees would not be particularly interested in an applicant who showed signs of altruism only in a high-stakes interview situation or in the period when the incentive to get into medical school provides the motivation for volunteering in various ways. However, most admission committees have little more to go on than an applicant's record of volunteering and his or her performance in an interview. Letters of recommendation may provide some additional information, but not necessarily. Further, because the letter writers are chosen by the applicant, it is difficult to know how much confidence to place in the information contained in the letters.

The admission committee must work within the confines of the information it has, but it is trying to project whether the applicant will always and foremost put the patients' and community's needs above his or her own when he or she becomes a physician. This selflessness manifests itself in the practicing physician in many different ways, from business relationships with referral partners, conflicts of interest from investments, responses to incentives provided by drug and equipment manufacturers, willingness to volunteer for free medical clinics, etc. The list is endless and the potential for conflicts of interest is similarly pervasive. In the admission process, the interest is not just in whether the applicant will succumb to one specific temptation, but whether he or she will choose altruism as a guiding principle in most situations, if not all. During an interview, for example, how does one separate an honest portrayal of an applicant's response to a hypothetical situation from a carefully crafted response the applicant thinks the interviewer wants to hear, or perhaps that the applicant has been coached to provide by a preparation service?

To what extent are these qualities influenced by nature, nurture, or maturation? Another issue pertains to whether personal qualities are truly stable. Do personality traits such as honesty truly solidify by age 18 as argued by Bullimore?[12] Are these truly "non-teachable" traits? If they are not teachable, are they malleable through immersion in a culture that values these traits? There seems to be considerable evidence that medical students change during medical school. Rezler,[26] in a literature review on medical students' attitude changes during medical school, stated that the medical school environment was largely responsible for the decreasing humanism in medical students and that curricular innovations are unlikely to result in more caring doctors until a majority of medical teachers model these necessary attitudes, skills, and dedication. Bland et al.,[27] in a comprehensive review of the literature on the determinants of primary care specialty choice, echoed Rezler's views on the decline in humanism during medical school and the negative influence of medical education. If medical school can have such a profoundly negative effect on students' humanism, it does not seem too far-fetched to suggest that

a properly focused medical school culture could promote positive personality characteristics. With a culture that values altruism and faculty who demonstrate altruism always and foremost, it is conceivable that students might be nurtured into becoming altruistic physicians. Further, if people mature at different rates, those at earlier stages of development may be even more likely to be affected by an altruistic culture.

The implications of this issue are particularly complex. If individuals do mature in their personal qualities such as altruism as they progress through medical school, the challenge for the admission process is to identify those who are most likely to mature in desirable ways. If personal qualities are stable, then the challenge for the admission process is to develop reliable and valid measures of these qualities and then to give them appropriate consideration. The problem may be that it is not an either/or proposition. Some personal qualities may be relatively malleable while others may be relatively stable by the time students enter medical school. To make matters even more complex, qualities that are malleable and those that are stable may vary among individuals. Sorting out this issue may be one of the greatest challenges to developing effective measures of personal qualities for medical school admissions.

What are the costs of measuring these qualities? The admission process is expensive for both the applicant and the institution, and adding measurements of new qualities in a rigorous manner will add more costs. Even if one uses information that is currently collected and derives new measures from it, or measures the qualities in a more rigorous manner, the change will add costs. The larger the number of personal qualities measured, the greater the costs incurred in their measurement. The new information must provide something of value beyond what was available before the new measures were added.

What are the ways of overcoming cunning adversaries? One of the greatest challenges facing any effort to systematically measure personal qualities will be the cunning ability of applicants and preparation services. Some applicants to medical school seem to have a chameleon-like ability to adopt the short-term personality of "Mother Theresa" and the career interest du jour. Further, the survival of some professional services depends on their ability to help applicants with repelling personal characteristics display compelling ones for a day. Sorting out fact from carefully crafted fiction will make developing a standardized measure of compelling personal characteristics a difficult challenge.

Some Possible Measurement Approaches

If schools attempt to take a least-cost approach to measuring compelling personal characteristics, making new uses of old information, (i.e. information already available) will be helpful. The American Medical College Application Service (AMCAS) application's Personal Statement and Essay could reflect compelling personal characteristics. Because all but a few medical schools already interview applicants, the interview itself might be a readily available source of information regarding compelling personal characteristics. Letters of recommendation are also commonly required to supplement the AMCAS application. Finally, there are elements of the transcript, parent(s)' education, and financial data that might give insights into personal qualities. For example, an applicant who is the first in his or her family to go to college, let alone go to medical school, is often given positive consideration in many medical schools. Other such insights might be gained from a careful and thoughtful consideration of such information.

Personal statement and essay. The personal statement is a relatively free-form essay that the applicant produces as part of the AMCAS application. A review of a sampling of such statements leads us to say that these statements are extremely variable. A literature search yielded no citation of its being used or evaluated to assess personal characteristics of the applicant. This would seem to be an untapped resource. To properly interpret the personal statement, it is important to know whether applicants believe the personal statement allows them to accurately represent their personal characteristics that qualify them for the profession of medicine, and whether the personal statement is a "group project" involving input from various others. To assess these issues, we surveyed matriculating first-year medical students for three years.[28] Across the three years, 53–84% of the respondents indicated that the personal statement adequately represented some element of their personal characteristics; 41–44% reported the personal statement involved input from others, with 15–51% reporting input in content development and 2–6% receiving input from professional services. Although the personal statement was considered by the large majority of matriculating medical students to adequately represent their personal characteristics, questions about help received in its preparation limit the confidence that admission committees can place in its accuracy for all applicants. Further, because of its free-form nature, any given personal statement will highlight a set of personal characteristics potentially different from the set highlighted in another applicant's personal statement. Making valid comparisons of applicants' personal characteristics from such non-standardized information offers significant challenges.

Interview. The interview is one of the few times, if any, prior to the admission decision when the applicant is physically present on the campus. Deciding how to use that precious time and to what purpose is a critical decision. The potential of the interview for assessing personal qualities extends beyond the two- or three-hour interview; the visit usually lasts at least twice that long and includes orientation

activities, tours, interactions with current students, luncheons, etc. Observations of applicants during these other times might also provide useful information.

Research suggests that structured interviews yield the most reliable estimates of personal qualities. A structured interview's features can include selecting interview content based upon a job analysis (e.g., the critical-incidents technique), standardizing the questions asked of all applicants, providing interviewers with sample answers to questions to help them give consistent ratings, and having the interview conducted by a board or panel of interviewers. One critical additional feature is providing training for the interviewers. Structured interviews require at least a minimal introduction to interview protocol and a rating system. Collins et al.[10] provided a half day of training for interviewers who were to be part of a 20-minute panel interview and then observed applicants interact in a group problem-solving session. Another potentially important issue concerns interviewers' qualifications. Patrick et al.[5] hired non-medical people to serve as interviewers and trained them in the protocol. Collins et al.[10] included representatives of consumer groups and experts in education along with medical school faculty on interview panels. Although non-faculty were used in these studies as interviewers, there was no effort to determine whether interviewers with different characteristics (consumer group representatives, education experts, medical school faculty) had distinguishable rating tendencies. Future research will need to determine to what degree interviewer characteristics produce detectable differences in interview results. There may also be an interaction between interviewer characteristics and applicant characteristics at play in the results.

Edwards et al.[3] considered a panel of interviewers to be part of the defining characteristics of a structured interview process. Having more than one interviewer enhances the reliability of the resulting ratings, but multiple interviewers can impact the dynamic of the interview in potentially unpredictable ways. Having multiple faculty interrogate an applicant can seem threatening because, even if the interview is structured to be collegial, the imbalance in numbers can be intimidating. In interviews conducted at our institution about our students' experiences in the admission interviews they had experienced at various schools, the students have been very critical of schools using a panel approach. Women students and students of color have been especially critical of schools using panel interviews. The interview format, interviewer characteristics, and applicant characteristics may represent a complex mix of factors that could have a major impact on the admission interview process.

A final concern relates to the nature of the interview. Collins et al.[10] reported two types of interviews being conducted: a 20-minute structured interview of one applicant by two panelists and two panelists observing six applicants as they participated in a 50-minute group exercise designed to stimulate debate. Separate panelists were used for the two types of interviews. Over 141 applicants, the correlation between ratings of the two interview types was .62. Although this is a relatively large correlation and was statistically significant beyond the .0001 level, it accounts for less that 40% of the rating variance. Disattenuating the correlation for the less-than-perfect reliability of the panel ratings (.67) still left over 50% of the variance in the two ratings unexplained, which clearly indicates that the nature of the activities and how they are structured for applicants may have a substantial effect on the results of the interview. The key is to structure the interview such that the personal qualities of interest can be assessed in a meaningful manner. Even if some personal qualities such as resourcefulness may be amenable to assessment in this type of situation, others such as altruism are likely to suffer from the artificiality of the conditions. Research related to how to measure various personal characteristics in an interview situation is clearly needed.

Much may be gained from the non-interview portion of the campus visit. The time applicants spend interacting with each other, participating in the orientation activities, meals, tours, etc. potentially can offer much insight into applicants' personal characteristics. One approach might be to adopt an element of the 360-degree evaluation model being explored for resident and physician evaluation by the Accreditation Council for Graduate Medical Education (ACGME) and American Board of Medical Specialties (ABMS). In this approach, almost everyone who comes into contact with the individual being rated provides a rating. In the admission interview case, one could do the same for the applicants during the non-structured interview time. Medical students, receptionists, food-service workers, tour guides, bus drivers, dean's staff, and others who interact with the applicants during the non-interview activities could be asked to rate the applicants. Because a relatively large number of applicants appear at the same time, such an evaluation would have to be picture-coded. It would probably be unreasonable for all of these different types of individuals to rate all applicants on all of the desired personal characteristics. Their contact would be so variable and transitory that it would probably be mostly wasted effort. If, however, these different individuals reported only memorable interactions of both positive and negative kinds, the strategy might provide useful information. At the very least, such an approach might be worth exploring.

One could also build tests of various personal qualities into the structured and unstructured portions of the visit. To reliably and validly measure personal qualities that are stable across time and situation will take care and creativity. The potential for building unobtrusive tests into the interview visit might be worth exploring, but care must be taken that it is done humanely and does not create a climate inducive of paranoia where the applicant feels under the microscope at every moment. As attempts are made to measure personal

characteristics, one needs to be mindful of the potential side effects that these measurements might produce.

Letters of support. Letters of recommendation are commonly requested as supplements to AMCAS application materials. Medical schools vary in how they instruct applicants to select letter writers and the degrees of structure imposed on what is to be written. The consequence is that letters of support are often difficult to interpret. Because the writers are chosen by the applicants and the formats are often relatively free-form, it is never clear how representative a given letter is of the applicant, and, if it is an accurate portrayal, how to evaluate the quality of one applicant's letters of support against those of another applicant.

Even more problematic is the risk of fraudulent letters of support. This is not just a theoretical possibility. We had to dismiss a student we had admitted to our medical school when it came to light that the letters of support were self-fabricated. Further, in focus-group meetings about the admission process, some students have admitted to participating in preparing some of the "letters of support."

Another challenge is that interpreting letters of support varies and often depends upon previous institutional experience with individual letter writers. A relatively bland letter from one writer might be considered extremely positive, and a comparable letter from a more effusive letter writer could be considered extremely negative. A more systematic approach to interpreting letters of support would be helpful. If national standards were to be developed about what should be included in letters of support for medical school, at least the content of the letters might be easier to evaluate. Further, there may be ways of applying content analysis procedures to letters of support that could aid in their interpretation. To address the concern about fraudulent letters, if a standard format could be adopted, perhaps an electronic system like that currently required by the National Institutes of Health (which has built-in security factors) could be used. These issues deserve further consideration and research.

Elements of the transcript, parent(s)' education, and financial data. Deducing compelling personal characteristics from the AMCAS file, supplemental application form, and various demographics is a complex task. Certain key markers are sometimes used to consider applicants for scrutiny, such as ethnicity, first generation to attend college, rural/inner city residence, low income, etc. The general purpose of interpreting these markers is to give applicants from underrepresented and disadvantaged backgrounds consideration for admission in the context of their backgrounds. Such analyses are challenging and often become politically charged issues. For example, an applicant who has overcome substantial adversity to make it to medical school may be barely keeping his or her head above water under the weight of life's demands, obligations, or lingering effects of earlier obstacles (single parenthood, emotional trauma from loss of relatives/ friends, surmounting a poor early educational system, etc.). The additional demands that medical school imposes can often push him or her under. Determining the difference between an applicant who can make it and one who cannot is difficult. In our experience, it often is not academic ability but the crush of life's obligations that makes the difference. This kind of issue is often discussed in the interview, but there may be ways to separate high risk from lesser risk through comparing the academic and demographic profiles of students who have made it through medical school with those who have not. More research in this area could yield valuable information.

A Proposal for A Unified System of Assessment

There are larger forces in the universe of medical education that might be usefully applied to assessment in the admission process. MSOP has identified four major objectives for medical education, each of which has six to 11 subobjectives. The ACGME and the ABMS have jointly identified six competencies for the practice of medicine. The MSOP competencies map fairly closely on these six competencies. Discussion of extending the ABMS/ACGME competencies into undergraduate medical education were held at the 2002 meeting of the Central Group on Educational Affairs, as was discussion of the possibility of integrating student-evaluation methods across undergraduate medical education (UME) and graduate medical education (GME).[29]

The potential benefits of integrating assessment methods across UME and GME would seem to be many. The ABMS and ACGME are beginning work on operationally defining how to measure the six competencies, beginning with communication skills. One of many tools that comprise the toolbox being developed for this purpose is the 360-degree evaluation. For a resident physician who is being evaluated, for example, one could have supervising physicians, nurses, patients, and administrative staff complete evaluations. Cost-effective methods of obtaining these evaluations are being developed.

Clearly, competencies appropriate for physician recertification would be more advanced than those for resident certification to practice, and, similarly, medical student competencies would be less advanced than those for residents. If we extend the concept of integration into the admission process, the competencies identified for applicants to medical school (pre-medical competencies) would be more rudimentary than those established for medical students. However, if one considers competencies to be a continuum from cradle to grave, the natural progression could serve as a means for assessing individuals at specific defining points. The evaluation methods used could build upon one another for continuity so that students feel a sense of progression and are better able to self-regulate their learning (monitor progress, identify learning needs, and adjust study accordingly). This Unified System of

Assessment would enable all parties interested in measuring the competencies of physicians and physicians-in-training to pool their resources and adopt a developmental approach to the measurement process.

Even now, some of the methods being developed by the ACGME/ABMS collaboration might have potential application to the admission process. As it progresses, the work of these groups may help to narrow the field of personal qualities that are of the highest priority for assessment. For example, the ACGME/ABMS collaboration has adopted the American Board of Internal Medicine (ABIM) process for peer and patient assessment. The ABIM recertification process involves having diplomates arrange for ten professional colleagues and 25 patients to answer ten questions about their overall medical care and communication skills.[30] They use a computer-administered telephone survey to collect the information. Using a similar method, it might be that a reasonably small number of applicant-nominated individuals completing a telephone survey can provide reliable and valid assessments of an applicant's personal qualities. If it is found that other personal characteristics can be better assessed via interview, the interview could be better focused to provide more reliable and valid measures of these other characteristics. It might even be possible to add a SAMS-type interview to the MCAT administration. This would potentially make assessing personal qualities during the campus visit optional or it could emphasize the elements unique to the particular school.

The segregation of UME, GME, CME and recertification has gone on for far too long. We need to consider the entire process as a continuum that includes even the selection of students for medical school and the pre-medical requirements. Pooling the resources of the entire system devoted to the education, testing, certification, and recertification of physicians would contribute to making much more headway than can be done with the current fragmented and separate small-scale efforts. Developing a continuum of competencies is a first step, developing a unified system for assessment would be the next.

REFERENCES

1. Cohen J. Facing the future. President's Address, 112th Annual Meeting of the Association of American Medical Colleges, November 4, 2001, Washington, DC.
2. Puryear JB, Lewis LA. Description of the interview process in selecting students for admission to U.S. medical schools. J Med Educ. 1981;56:881–5.
3. Edwards JC, Johnson EK, Molidor JB. The interview in the admission process. Acad Med. 1990;65:167–77.
4. Nowacek GA, Bailey BA, Sturgill BC. Influence of the interview on the evaluation of applicants to medical school. Acad Med. 1996;71:1093–5.
5. Patrick LE, Altmaier EM, Kuperman S, Ugolini K. A structured interview for medical school admission, Phase 1: initial procedures and results. Acad Med. 2001;76:66–71.
6. Meridith KE, Dunlap MR, Baker HH. Subjective and objective admissions factors as predictors of clinical clerkship performance. J Med Educ. 1982;57:743–51.
7. Murden R, Galloway GM, Reid JC, Colwill JM. Academic and personal predictors of clinical success in medical school. J Med Educ. 1978; 53:711–9.
8. Powis DA, Neame RLB, Bristow T, Murphy LB. The objective structured interview for medical student selection. BMJ. 1988;296:765–8.
9. Taylor TT. The interview: one more life. Acad Med. 1990;65:177–8.
10. Collins JP, White GR, Petrie KJ, Willoughby EW. A structured panel interview and group exercise in the selection of medical students. Med Educ. 1995;29:332–6.
11. Shaw DL, Martz DM, Lancaster CJ, Sade RM. Influence of medical school applicants' demographic and cognitive characteristics on interviewers' ratings of noncognitive traits. Acad Med. 1995;70:532–6.
12. Bullimore DW. Selection interviewing for medical school admission. Med Educ. 1992;26:347–9.
13. Richards P, McManus IC, Maitlis SA. Reliability of interviewing in medical student selection. BMJ. 1988;296:1520–1.
14. Van Susteren T, Suter E, Romrell LJ, Lanier L, Hatch RL. Do interviews really play an important role in the medical school selection decision? Teach Learn Med. 1999;11:66–74.
15. Litton-Hawes E, MacLean IC, Hines MH. An analysis of the communication process in the medical admissions interview. J Med Educ. 1976;51:332–4.
16. Smith SR, Vivier PM, Blain ALB. A comparison of the first-year medical school performances of students admitted with and without interviews. J Med Educ. 1986;61:404–6.
17. Balogunn JA. Predictors of academic and clinical performance in a baccalaureate physical therapy program. Phys Ther. 1988;68:238–42.
18. Walker JD, Killip DE, Puller JL. Significance of the admission interview in predicting students' performance in dental school. J Med Educ. 1985; 60:569–71.
19. Tutton P. The selection interview. J Higher Educ. 1997;19:27–33.
20. Mitchell KJ. Traditional predictors of performance in medical school. Acad Med. 1990;65:149–58.
21. Vancouver JB, Reinhart MA, Solomon DJ, Haf JJ. Testing for validity and bias in the use of GPA and the MCAT in the selection of medical school students. Acad Med. 1990;65:694–6.
22. Albanese MA, Farrell P, Barnet J. Sensitivity and specificity of using GPA and MCAT cutoffs for medical school admissions decisions. Proceedings of the 38th Annual Research in Medical Education Conference, November 1999, Washington, DC.
23. Price PB, Lewis EG, Loughmiller GC, Nelson DE, Murray SL, Taylor CW. Attributes of a good practising physician. J Med Educ. 1971;46:229–37.
24. McGaghie WC. Assessing readiness for medical education: evolution of the Medical College Admission Test. JAMA. 2002;288;1085–90.
25. The Medical School Objectives Writing Group. Learning objectives for medical student education—guidelines for medical schools: report i of the Medical School Objectives Project. Acad Med. 1999;74:13–8.
26. Rezler AG. Attitude changes during medical school: a review of the literature. J Med Educ. 1974;49:1023–30.
27. Bland CJ, Meurer LN, Maldonado G. Determinants of primary care specialty choice: a non-statistical meta-analysis of the literature. Acad Med. 1995;70:620–41.
28. Albanese MA, Snow M, Skochelak S, Huggett K, Farrell PM. Assessing personal qualities in medical school admissions: matriculate perceptions of the contributions of the interview and personal statement. Presented at the 2002 meeting of the Central Group on Educational Affairs, Chicago, IL, March 2002.
29. Anderson B, e-mail communication to Mark Albanese, July 25, 2002.
30. Wasserman SI, Kimball HR, Duffy FD. Recertification in internal medicine: a program of continuous professional development. Ann Intern Med. 2000;133:202–8.