Noncognitive Factors That Predict Professional Performance

Obviously, three requirements are fundamental: mental equipment, physical equipment, and that quality so difficult to define—character.

*John Wyckoff, 1927*

We are going to be forced to pay more attention to other qualities of character and personality. . . . These things are hard to define and difficult to measure, but they may be the most important factors.

*Philip B. Price, 1964*

In the previous chapter I traced the origins of the current system of evaluating applicants to medical school based largely on their cognitive abilities, measured as success in the undergraduate premedical sciences. However, beginning more than eighty years ago, medical educators such as those cited above were voicing concerns about including in the evaluation process assessment of noncognitive characteristics such as “character” and “personality.” Addressing the Association of American Medical Colleges (AAMC) in 1930, Dr. Edward Thorpe of the University of Pennsylvania Medical School cautioned, “If more courses in chemistry, biology, and physics are needed to prepare students for the medical school curricula of the present and future, what effect has maximum preparation had on students in the past? . . . I have noted with some misgivings the depreciation in that abstract quality called culture in the average applicant for medical training.”¹ At that same meeting, E. P. Lyons, dean of the University of Minnesota Medical School, delineated what was meant by the term *culture*: “Culture is the antithesis of specialism, of research, of scholarship. Culture is broad; to the extent that our medical curriculum is narrowing, it is anticultural.”²

Between 1927, when Wyckoff made his comments quoted in the epigraph as part of his research on how to select students for admission to medical school,³ and 1956,
when the AAMC held its training institute titled “The Appraisal of Applicants to Medical School,” medical education researchers paid scant attention to including assessment of an applicant’s “character” or “culture” as part of the medical school admissions process. As described in the previous chapter, the AAMC’s training institute was followed by a shift in the focus of research on predicting the outcomes of medical education. The discussions at the institute were equally as concerned with considering an applicant’s intellectual, or cognitive characteristics as they were with considering the non-intellectual, or noncognitive characteristics of applicants. The latter discussion concluded, “It is evident that the so-called nonintellectual characteristics of medical applicants—the elusive personal and social traits and motivational factors—share importance with the more readily measurable intellectual traits in the appraisal of applicants to medical school.”

How should one measure in a reliable and valid manner the noncognitive characteristics of an individual interested in a medical career? This question absorbed much of the attention of the conference attendees. The simplest way to address this question was to include a personal interview with the applicant as part of the admissions process—a practice that was becoming increasingly prevalent among medical schools at that time. While the institute participants were in general agreement on the importance of the admissions interview, they nonetheless acknowledged some skepticism about the validity of the interview in evaluating noncognitive attributes: “As long as we have no objective way of determining our criteria for good physicians of whatever variety, and as long as we have no objective ways of evaluating some of the traits that are allegedly prerequisites for becoming a good physician, the interview is the only tool we have for estimating traits.” The report of the meeting called for additional research to confirm or support the role of the interview in assessing the noncognitive characteristics of applicants to medical school. This chapter will consider the results of this research.

In addition to using a personal interview to evaluate applicants, two other approaches were discussed at the conference: psychiatric interviews and standardized psychological tests. Dr. Joel Handler reported on the results of psychiatric interviews conducted with more than 500 medical students from three entering classes at the University of Illinois College of Medicine. Dr. Handler and his colleagues had not actually used the results of psychiatric interviews in the admissions process. Rather, they had administered these interviews to incoming students on a voluntary basis (although only two of the more than 500 incoming students declined to be interviewed). Dr. Chandler’s findings were not particularly flattering regarding the psychiatric profile of the incoming students: “Our impression was that the majority of the students we saw were quite conformist, emotionally constricted men and
women, given to internalizing their hostility either with depressive or compulsive traits for the most part.” Chandler’s description of the students he interviewed are certainly not what Dr. Wyckoff had in mind when he described “that quality so difficult to define—character.”

Regardless of what Dr. Chandler found, it seemed neither practical nor ethical to use a routine psychiatric assessment as a tool with which to evaluate applicants to medical school. However, the use of standardized psychological assessments presented the opportunity to get at many of the characteristics Dr. Chandler was addressing. There was extensive discussion at the AAMC’s 1956 meeting of the potential role of psychological tests in improving the admissions process for medical schools. Subsequent to this meeting, the AAMC began its Longitudinal Study, in which researchers administered a battery of psychological tests to entering medical students at 28 different medical schools and then followed these students into their professional careers. Results from this study were in two papers presented to the annual meeting of the AAMC in 1963.

In the first paper, Charles Schumacher of the National Board of Medical Examiners described the purpose of the AAMC’s Longitudinal Study as providing an answer to two questions: (1) What are the psychometric tools that will permit medical schools to improve their technique for selecting medical students? and (2) What are the psychometric tools with which medical schools might improve their techniques for counseling and guidance of medical students? Schumacher looked at five different specialty areas in medical practice (general practice, internal medicine, surgery, medical research/teaching, surgical research/teaching), comparing individuals within these groups on three factors: scholastic aptitude, personality characteristics, and biographical histories. He found the specialists engaged in teaching showed higher scholastic aptitude, as measured by their MCAT scores, and personality characteristics consistent with the desire to create new knowledge. By contrast, the practicing physicians, especially the general practitioners, scored lower on the MCAT while showing personality traits that were conducive to “the practical application of knowledge to the everyday problems of patient care, rather than in the creation of new knowledge.” To Schumacher it was clear that, upon entering medical school, students displayed personality traits as measured on standardized psychological tests that could be used to predict the type of medical career they would select. Schumacher made no attempt to suggest that one set of personality traits was preferable to another. Rather, he simply confirmed that traits differed in predictable ways among students headed in different career directions.

In the second paper presented at the 1963 meeting, Edwin Hutchins, a researcher from the AAMC, compared the psychological profiles of entering medical students.
with the profiles of graduating college students in general. He found a series of differences, including the following:

- Medical students scored higher on theoretical and aesthetic scales, suggesting suitability for intellectual pursuits.
- Medical students tended to score higher on scales of individualism and self-sufficiency.
- Medical students scored lower on scales of social and altruistic values.
- Medical students tended to be more conventional, with less need for independence in decision making.\(^\text{11}\)

For nearly three decades John Caughey served as associate dean for admissions at Western Reserve University School of Medicine (now Case-Western Reserve). In this capacity he was actively involved in AAMC activities and for a number of years served as chairman of the AAMC’s Group on Student Affairs, which had been established as a result of the discussions at the 1956 training institute. In 1966 he reported to the group on the ten-year follow-up of the work on the psychological assessment of medical students and medical school applicants that had been initiated at the 1956 institute. He commented that medical educators continued to do a good job of assuring an adequate level of scientific knowledge of all medical students. However, he was concerned about the relatively little weight admissions committees gave to the “non-intellectual components of medical education, which are of critical importance in the development of the physician.”\(^\text{12}\) Caughey stressed the importance of personal integrity as a characteristic of at least equal importance as premedical academic attainment in identifying students qualified for the study of medicine. However, medical schools continued to function without a reliable means of assessing integrity. Caughey was also concerned that, on realizing the emphasis medical schools placed on scientific knowledge over noncognitive personal qualities, premedical students would themselves de-emphasize improving those very noncognitive qualities.

In response to the results of the AAMC’s Longitudinal Study reported above by Hutchins and with direct reference to Caughey’s remarks, Parlow and Rothman reported on an effort at the University of Toronto Faculty of Medicine to incorporate and give added weight in the admissions process to measures of noncognitive characteristics. To do this, they administered a series of psychological tests to incoming medical students over the period 1967–72. They found that by incorporating this dual emphasis in the admissions process they were able to select those students who scored highly on both those cognitive characteristics conducive to academic success and those noncognitive characteristics, such as “nurturance,” that are, “the most germane to medical practice.”\(^\text{13}\)
Haley and Lerner undertook similar research in which they administered a battery of psychological tests to 114 medical students in two successive entering classes at the University of Kentucky College of Medicine. They then followed the students through the first two years of medical school to look for both cognitive (MCAT scores and undergraduate grades) and noncognitive (psychological traits) predictors of academic success. Their findings raise some critical issues: “In general the results indicate that students who do well in the majority of courses . . . tend to have submissive and uncritical attitudes toward authority and a relatively cynical view of human conduct. They are comparatively ambitious to achieve personal, political, or economic power . . . and less socially concerned than students who do less well in these courses.”14 In looking to see which psychological traits were associated with success in the pre-clinical science courses, the authors found that “terms [such] as intelligent, sensitive, or socially concerned . . . either had no association or were negatively associated with the majority of courses.”15

Turner and colleagues reported on their research at the College of Medicine at Ohio State University in which they administered a series of psychological tests to 50 incoming medical students and then compared the results of these tests with students’ performance in a videotaped interaction with a patient.16 In much the same way that the Standardized Patient Examination would be scored thirty years later, students were rated in three areas of clinical quality: communication skill, interpersonal skill, and physical examination skill. While they found significant positive associations between measured clinical skills and initial psychological measures such as “emotionally stable” and “judgment-perception,” they found a significant negative association between clinical skills and MCAT science subtest scores. These results led the authors to encourage the use of psychological tests more broadly in medical schools.

Researchers from the University of California, Berkeley, administered psychological tests on an experimental basis to incoming medical students at the University of California, San Francisco (UCSF), between 1955 and 1967. They then followed these students through medical school to address two questions: (1) Did the results of the psychological tests predict academic performance in medical school? and (2) Did the results of the psychological tests predict failure to graduate? Consistent with previous research on cognitive performance, Gough reported that MCAT science scores and undergraduate science GPA were associated with grades in the first two years of medical school at UCSF, but were “almost completely unrelated to performance in the fourth year and to faculty rating of general and clinical competence.” However, those same students who excelled in science as premedical students were found, on the psychological tests, “to be narrower in interests, less
adaptable, less articulate, and less comfortable in interpersonal relationships than their lower scoring peers.”

In their study of the factors associated with failure to graduate, Gough and Hall found that, of 1,071 incoming medical students at UCSF, 1.6 percent withdrew from school for academic reasons and 3.8 percent withdrew for personal, nonacademic reasons, for an overall failure rate of 5.4 percent. They found that the psychological tests administered at the beginning of school were somewhat better at predicting failure than were the traditional predictors of MCAT scores and premedical grades. Based on this research, they suggested that admissions committees use a combination of cognitive and noncognitive assessments to get a better picture of those students most at risk of failure during medical school.

The results of Gough’s research are of particular interest to me because I entered the UCSF medical school as a first-year student in the fall of 1968, the year immediately following Gough’s twelve-year study. As a student activist, I was permitted to become the first medical student ever to sit on the UCSF medical school admissions committee, a seat I held from 1968 to 1973. While Gough’s work was not published until two years after I left the admissions committee, not once did I hear his results discussed within the committee. The issue of using as part of the admissions process an applicant’s noncognitive characteristics as measured on standardized psychological tests was never broached. As far as we were concerned, the principal characteristics qualifying an applicant for admission were high MCAT science scores and high grades in the premedical sciences. The personal interview was useful to flush out more information about an applicant’s personality, but it was not used in a standardized way. At least at UCSF at the end of the 1960s, the admissions process paid little attention to the recommendations stemming from the AAMC’s 1956 workshop or to Gough’s own conclusions that the admissions practices used at that time by UCSF and other medical schools “may be overemphasizing scientific talent and interests and thereby underemphasizing other desirable talents and attributes.”

A review published in 1989 by Miller and colleagues confirmed that slightly more than half of U.S. medical schools incorporated consideration of noncognitive characteristics into their admissions process but that they did so mostly by following general statements about considering characteristics such as a candidate’s honesty, dedication, or consideration for others, without any formal assessment of these traits.

While U.S. medical schools failed to incorporate the standardized assessment of psychological characteristics into their admissions process, at least one school outside the United States did so based on the growing scientific support for such an approach. In 1985, Feletti and colleagues reported on their experience at the Newcastle Medical School in New South Wales, Australia. Students in Australia were ac-
cepted into medical school after finishing secondary school in a system of admissions similar to the historic European model. The traditional means of selection was a student’s score on a series of standardized academic aptitude examinations. Acknowledging that doubts had been raised over the use of this admissions process, the admissions committee adopted a policy of admitting half of the entering class based on the traditional measures of academic attainment and half of the class based on the results of a series of psychological assessments of students’ strengths in areas such as empathy, creativity, response to moral dilemmas, and skills in abstract reasoning (though students admitted based on their psychological characteristics were required also to achieve a certain threshold level of academic attainment). The authors’ general conclusion was that the students who were selected based on their psychological profile were equally as successful in medical school as their classmates whose selection was based purely on academic attainment.

Despite the failure of most medical schools to adopt formal psychological assessment as part of the admissions process, a number of researchers in the 1980s and 1990s continued to study the associations between psychological characteristics and medical school performance. For example, in 1987 Aldrich reported a follow-up study of the use of formal psychiatric assessment of incoming medical students, suggesting that such interviews could identify students exhibiting passivity, social isolation, or unusual levels of anxiety or rigidity, and that such students might be at increased risk of failing to graduate for personal rather than academic reasons.22

Hojat and colleagues published two papers describing their administration of a series of standardized psychological tests to second-year medical students, then following those students through medical school and into their residency training. Their first analysis compared the use of MCAT scores with the use of the results of the psychological tests to predict grades on the basic sciences examinations in the first two years of medical school, grades on the clinical science examinations in the six core clerkships, and ratings of clinical competence given by the instructor in each of the six core clerkships. The authors found that “the admissions measures and the psychological measures predicted the students’ basic sciences and clinical sciences examination grades equally well, but the psychological measures were better predictors of the students’ clinical ratings than were the admission measures.”23 When the authors followed the students into their residency, they again found “significant links between selected psychosocial measures and physician clinical competence ratings.” After placing their findings in the context of their earlier research and that of others, they suggested, “It is not only mental abilities but also psychological attributes or emotional quotient (EQ) that enhance professional effectiveness above and beyond the well-known concept of intelligence quotient (IQ).”24
In 1990 William McGaghie of the University of North Carolina School of Medicine published a two-part review of an extensive literature on the medical school admissions process, paying particular attention to the use of what he called “qualitative variables.” 

Consistent with the other research described in this and previous chapters, McGaghie concluded that there was at best a weak link between a student’s academic performance as an undergraduate and his or her achievement across the full medical school experience. Furthermore, he cited considerable evidence that, while medical schools pay what he referred to as “lip service” to the importance of noncognitive attributes such as character, motivation, and personality, they continued to select students for admission based on high MCAT scores and high grades in the premedical sciences: “Despite widespread acknowledgement that qualitative factors are crucial for success as a medical student and physician, the variables are rarely measured or considered when medical schools reach decisions about student admission.”

McGaghie concluded his analysis by suggesting that, since the time in the 1960s when medical educators first became aware of the value of assessing noncognitive characteristics, there had been little if any progress in actually incorporating formal evaluation of these qualities into the admissions process.

In the last several years, we have continued to see research reports on the importance of noncognitive factors in predicting success in medical school, nearly all consistent with the research that preceded them. Shen and Comrey looked at the medical school experience of 97 economically or socially disadvantaged students who had attended the UCLA school of medicine, confirming that the traditional measures of undergraduate academic performance were associated with academic success in the early years of medical school, while personality characteristics were more predictive of success in the clinical years.

Cariaga-Lo and colleagues confirmed that most medical school delays or failures occur in the first two years and that a combination of cognitive and noncognitive measures is best suited to identifying those students who may experience delay or difficulty. Carrothers and colleagues reported testing a new multi-item measure of what they referred to as “emotional intelligence,” which, they concluded, “is weakly correlated with indicators of factual knowledge and analytic skills, and strongly correlated with the traditional instrument that measures how well a student will perform the social role of being a physician.” Finally, Manuel and colleagues tested whether personality factors were associated with the clinical skills of second year medical students, and were able to confirm such an association.

Success in medical school involves academic success. This has been clear since the 1920s or before. Success in medical school, however, also involves the development and expression of aptitudes that go considerably beyond scholastic aptitude. Only
when a student is able to exhibit the full range of requisite aptitudes is she or he fully prepared to become a physician. In 1996 the AAMC initiated a Medical School Objectives Project (MSOP) and charged its participants with identifying those attributes that all medical students should have by the time they have completed training. In 1999 the MSOP Writing Group issued its report, identifying the core attributes students will develop in order to consider their medical education to be a success:

1. *Physicians must be altruistic.* They must be compassionate and empathetic in caring for patients, and must be trustworthy and truthful in all their professional dealings.
2. *Physicians must be knowledgeable.* They must understand the scientific basis of medicine and be able to apply that understanding to the practice of medicine.
3. *Physicians must be skillful.* They must be highly skilled in providing care to individual patients.
4. *Physicians must be dutiful.* They must feel obliged to collaborate with other health professionals and to use systematic approaches for promoting, maintaining, and improving the health of individuals and populations.\(^{31}\)

The delineation of these requisite attributes provides medical schools with a means of evaluating their own curricula and gauging their own success. While it suggests a framework that might be applied to the selection of students for medical school from the applicant pool, most schools have yet to incorporate formal assessment of the noncognitive characteristics that have been shown to be associated with the full development of these attributes. As early as 1976 a Noncognitive Working Group established by the AAMC had recommended including assessment of a series of noncognitive characteristics in the MCAT. The project was never implemented.\(^{32}\) In a review of the barriers to incorporating formal assessment of noncognitive characteristics in the admissions process, Albanese and colleagues identified three such barriers: “institutional self-interest, inertia, and philosophical and historical factors.”\(^{33}\)

Medical educators in the United States are not alone in being reluctant to place added emphasis on the assessment of noncognitive characteristics. Reviewing medical education in the United Kingdom, Ferguson and colleagues questioned the traditional emphasis on scholastic attainment and the de-emphasis of noncognitive characteristics: “Previous academic ability, personal statements, references, and interviews are all traditionally used in selection, but how good are they at predicting future performance? Personality and learning style are not traditionally used, but should they be?”\(^{34}\)

The model of basing admission to medical school on scholastic aptitude, princi-
ally in the sciences, evolved in the United States over a period of several decades in an attempt to incorporate the European model of medical education into our own. The growing body of evidence that success in medical education and in medical practice depends on a substantially broader range of factors than scholastic aptitude alone offers us the opportunity to reassess the methods we use to select from among applicants for admission those students best suited to the future practice of medicine. Finding ways to accomplish this goal continues to present substantial challenges, but not necessarily challenges that cannot be overcome.

The Admissions Interview as a Means to Assess the Noncognitive Strengths of Applicants

The AAMC’s four-day meeting in 1956 to address “The Appraisal of Applicants to Medical School” looked separately at the appraisal of intellectual characteristics and of “nonintellectual” characteristics. Participants discussed two principal means of evaluating the noncognitive characteristics of medical school applicants: formal psychological tests and admissions interviews. As we have discussed, with rare exceptions psychological tests have not been incorporated into the admissions process in the United States. The admissions interview, however, has been part of that process for several decades. In 1956 all of the 89 medical schools responding to the AAMC’s survey about the sources they use for evaluating noncognitive characteristics indicated that they used the personal admissions interview, with 72 percent of the respondents indicating that they relied heavily on the interview in this regard.

Despite the universal use of the admissions interview, the introduction to the report in which these data were presented explicitly acknowledged the dilemma facing the schools: “It is clearly evident that some of the characteristics valued most highly by medical schools are those which they have least confidence in evaluating. . . . For the nonintellectual characteristics, there appears to be a negligible correlation between the recognized importance of these factors and the degree of confidence in their evaluation.”35 There was consensus among participants at the meeting that the personal interview had little role in predicting the academic success of medical students in the early years of medical school. Rather, the interview served principally to clarify information in the formal application form and permit the applicant to explain any unusual aspects of the application, to screen for what were referred to as “gross deficiencies in personality and emotional stability,” and to gather personal information about the applicant that can be used at the margins of the decision process when there are more applicants who are academically qualified than there are available admissions slots.36
Despite acknowledging the importance of predicting success in both the preclinical sciences of medical school and the clinical aspects of medical school, participants were reluctant to use the interview to predict future clinical success for a simple reason: “Before one can test whether the interview can be used to predict future performance of medical applicants, it will be necessary to establish criteria of success in medicine.”

Participants in the 1956 meeting held a symposium focused exclusively on “The Interview as One Tool for Selection,” with the presentation and discussion of a series of papers. Joseph Zubin, a biometrician, suggested that the purpose of the interview “is to gather information about the candidate’s motives, feelings, attitudes, and integrity insofar as they determine his interests in medicine and his ability to deal with peoples.” However, Zubin also pointed out some potential problems in using the interview in this manner, principal of which is the “halo effect” by which the interviewer forms a general impression of the interviewee based on information contained in the formal application (e.g., grades, MCAT scores) and allows that general impression to influence his or her assessment of the applicant’s noncognitive characteristics as well.

Zubin’s paper was followed by one presented by E. Lowell Kelly, a professor of psychology from the University of Michigan. Kelley described his own research and that of others demonstrating that there simply was no evidence of the validity of the interview as a predictor of future performance, either in medical school or in medical practice. Medical schools believed strongly in the utility of the interview, and the longer they used the interview as part of the admissions process the more they tended to believe in them. That belief, however, was not supported by scientific evidence.

Kelley also presented a largely unrelated but nonetheless interesting result from his research on the performance of medical students at the University of Michigan. In evaluating grades throughout the medical school experience, clinical as well as preclinical, he found that “the number of premedical credit hours in inorganic chemistry and in biology submitted by the applicant is negatively correlated with medical school grades.” Kelly was one of the first researchers to suggest dual predictive associations for the amount of science a student had included in his or her premedical curriculum: more science predicted better performance in the preclinical sciences of medical school but worse performance in the clinical aspects. We will see this inverse association again in the discussion of the role of empathy in medical school success.

The 1956 meeting called repeatedly for more research on the use of the interview in the medical school admission process. In 1990 Edwards and colleagues published
a review of that research. They were able to identify several important trends and results about the role of the interview. While acknowledging the “halo effect” as a potential pitfall, they suggested that the use of structured interviews with training for the interviewers offered a way to avoid that effect. When interviews were appropriately structured, there was “some evidence, albeit imprecise, that the interview actually predicts clinical performance in medical school.” To do this best, the interview should de-emphasize information pertaining to academic performance and focus instead on identifying personal characteristics of the applicant such as “leadership, motivation, range of interests, and interpersonal skills.” The key to success in this regard is for each medical school to decide for itself what constitutes “success” in medical practice and to look for those noncognitive characteristics that have been shown to be associated with success measured in that way.41

In an editorial accompanying the paper by Edwards and colleagues, Thomas Taylor of the University of Iowa College of Medicine expressed skepticism about the role of the interview: “The interview is well entrenched in the admissions process, and it has the validity that comes from habit. Everyone is used to it. It is like an old cat. And like that old cat, it probably will hang around for a while, though nobody can really explain why.”42

Harasym and colleagues conducted research in which they included some “standardized applicants” in the interview pool at one Canadian medical school. Similar to the standardized patients used as part of medical licensure examinations, the standardized applicants were trained actors enacting scripted roles. Each actor was interviewed by six different interviewers, and the scoring of the interviewers was compared to assess the statistical reliability of the scoring process. The results were not especially encouraging. The researchers found significant variability among the scores of the various interviewers. They did note, however, that the more experienced the interviewer, the more consistent was the scoring.43

Despite questions about the reliability of interview scores, Elam and Johnson, in a study of voting patterns in a medical school admissions committee, noted that interview scores had a strong and direct association with the number of supporting votes cast by committee members for an applicant.44 Also looking at admissions committee voting patterns, Georgesen and colleagues addressed a somewhat different question: for those applicants deemed acceptable but placed on an alternate list pending the results of the first round of offers of admission, what factors are associated with a candidate’s position on the alternate list? As admissions officers go down the alternate list as slots open up, an applicant’s position on that list can be a key determinant of his or her chances of subsequently being admitted. They found that an applicant’s interview scores were the strongest predictor of position on the list, but
in a reverse manner. Each applicant had two interview scores. The higher of the two scores had little association with placement on the list, while the lower score had a strong association with placement. Admissions committee members were hesitant to give priority to an applicant with a single low interview score, while paying relatively little attention to a conflicting but higher score.\textsuperscript{45}

In sum, it seems that, if done well, an admissions interview can elicit information regarding noncognitive characteristics that is useful in predicting future academic and professional success. To be done well, an interview should be consistently structured so as to elicit the same information across candidates. The interview should be designed to assess those specific noncognitive characteristics thought to be important by the medical school. Interviewers should be trained in its use. When done in this manner, the interview gathers important information that will have a low correlation with traditional measures of cognitive ability, such as premedical sciences grades and MCAT scores, as confirmed by the research of Patrick and colleagues.\textsuperscript{46}

The Role of Empathy in Medical Care and in the Selection of Medical Students

Deciding which noncognitive characteristics to assess as part of the admissions process depends on how an admissions committee views the concept of professional success in medicine. The AAMC addressed this issue in a report published in 1984 titled “Physicians for the Twenty-first Century.” In the report, a panel of researchers and educators affirmed that “all physicians, regardless of specialty, require a common foundation of knowledge, skills, values, and attitudes.” The panel identified one specific attribute as primary: “We believe that every physician should be caring, compassionate, and dedicated to patients.”\textsuperscript{47} Caring and compassion require empathy—a sense of emotional understanding and connectedness with the patient.

Empathy as an emotional reaction to an interaction or observation has been recognized for more than two centuries as distinct from a cognitive or intellectual reaction to the same situation. Empathy is a concept that can be measured along multiple dimensions, such as concern, adopting another’s perspective, and sensing another’s distress. Psychologists have been able to develop reliable methods of measuring empathy and have found that it is a characteristic largely independent of scholastic attainment. When researchers in one study administered psychological tests to college students in an introductory psychology course, they found that the various measures of empathy had little if any correlation with measures of cognitive ability such as the subscales of the Scholastic Aptitude Test.\textsuperscript{48}
In 1978 David Kupfer, a faculty psychiatrist, and Frances Drew, the student affairs dean, both at the University of Pittsburgh School of Medicine, published a provocative article regarding the personality characteristics of medical students. The authors acknowledged the growing body of research at the time suggesting that MCAT scores and grades in the premedical sciences, while having substantial power to predict academic performance early in medical school, had little power to predict performance in the clinical years. In their own school they had administered a standardized instrument measuring empathy to more than 500 students and then looked for correlations between MCAT scores and empathy scores. They found few significant relationships, with those that did reach statistical significance being fairly small in magnitude. They concluded that empathy and other similar personality characteristics reflect dimensions that are separate and distinct from cognitive ability. Given the importance of empathy and the large number of applicants who have the academic strength to succeed in medical school, they suggested that medical schools look first for those applicants who are high on both dimensions: cognitive ability and empathy.49

Branch and colleagues describe medical students’ own perspectives on empathy, using students’ written narratives of their interactions with patients. In many of these narratives, students acknowledged the critical importance of connecting with their patients empathetically: “To a remarkable degree, the medical students put themselves in their patients’ shoes... the students’ learning to understand patients’ experiences through empathy often improved their relationships with patients.”50 There is another thing to be learned from reading these narratives. Not only do students recognize the value of their own ability to empathize with patients, they also readily recognize a lack of empathy or compassion on the part of their student colleagues or the practicing physicians with whom they work and the adverse effects of such a lack.

Empathy is “the ability to understand another person’s emotional or life experience; it is to share those emotions’ content but not their intensity.” Using this definition, E. R. Marcus links empathy with the distinct but related quality of humanism and identifies them as capacities that are critical to the practice of medicine. From Marcus’s perspective, empathy reflects an ability, and humanism an attitude. Unfortunately, based on Marcus’s experience training medical students and residents in psychiatry, the experience of medical school often is inimical to the development of empathetic ability and a humanistic attitude: “Constant picayune testing, a threatening academic atmosphere, and a competitive curve grading system increase student competition and student anxiety, and reinforce the grandiose ‘must master it all’ defense. The bigger this defense, the bigger the inevitable crash and the
greater the resultant humiliation and anger. The result is a tough emotional crust and marked disidentification with patients: in distancing themselves from their own victimization by the curriculum, students distance themselves from the victims of illness.” Marcus suggests that what he refers to as “the rites of passage” of medical school delay the final emotional maturation of the physician until after he or she has completed residency training. This maturation process takes place unconnected to any formal educational process and is subject to variation among individuals.51

Whatever the level of empathy with which students enter medical school, research suggests that level will decline over the course of medical education. Diseker and Michielutte assessed students’ empathy during the first week of medical school and at the end of the fourth year. Using an empathy scale derived from the California Personality Inventory (CPI), a well-established and previously validated personality assessment, they found a decline in the average level of empathy over the four-year period. In their study, empathy correlated negatively with MCAT scores.52

West and colleagues, using a different measurement instrument, evaluated the level of empathy and the level of medical knowledge of entering residents in an internal medicine residency program. They then repeated these assessments at the beginning of the second year. Over the period of one year, these residents demonstrated an increase in their level of medical knowledge but a decrease in their level of empathy. In discussing the implications of their results for the broader process of medical education, the authors concluded “that the core competencies [medical knowledge and empathy] represent separate domains of skill/expertise that develop independently . . . our findings confirm the importance of measuring each domain of competency separately.”53

Might the rites of passage inherent to medical education, identified above by Marcus, actually begin well before medical school, in the early parts of premedical education? “Constant picayune testing, a threatening academic atmosphere, and a competitive curve grading system,” the factors described above by Marcus as impeding the development of empathy, also characterize the early premedical experience, as our interviews with premedical students at Stanford and UC Berkeley, described in the first chapter, so clearly indicate. The academic and personal strengths necessary to succeed in a highly competitive premedical curriculum might be contrary to those strengths that support the development of empathetic ability. Could it be that success in the premedical and preclinical sciences is negatively associated with empathy? This question was addressed by Peter Tutton, a member of the medical faculty at Monash University in Australia.54

The medical school at Monash University has a six-year curriculum similar to the historic European model of medical education, with students entering after high
school. During the first three years, incoming students study the premedical sciences, with completion of the preclinical sciences and clinical training in the final three years. Tutton administered the CPI to 133 incoming students. He then followed the students through their first three years, comparing their grades in the sciences to their scores on the various scales of the CPI. He found a consistent pattern of correlations.

The principal type of examinations used in the sciences at Monash University involves multiple choice questions (MCQs). Success in MCQs in the sciences was significantly correlated with three of the various personality scales measured by the CPI: empathy, dominance, and internality. The correlation with empathy was a negative one, with a correlation coefficient of $-0.36$. The better students did on the their science exams through their first three years of college, the lower was their empathy score on entering college. Conversely, those students with the highest empathy as measured by the CPI tended to do worst in their science examinations. The correlation with the scale of dominance was also negative ($-0.32$), while that with the scale of internality was positive ($+0.30$).

As explained by the author, a high score on the internality scale, as students who were successful in their MCQs tended to get, suggests that a student may be “shy,” “submissive,” “withdrawn,” or “awkward and ill at ease socially,” characteristics the author suggests are “the antithesis of what most of us would want in a clinician.” Summarizing his results, Tutton further suggests that “students with high achievement in many components of the curriculum tend to have personality profiles that seem inappropriate for their chosen careers as physicians.”

For more than a century, success in the premedical sciences of chemistry, biology, and physics has been seen as a necessary precursor to success in medical school and ultimate success as a physician. Since the 1920s success in the premedical sciences, as measured by undergraduate grades and MCAT scores, has been the principal means of sorting medical school applicants to identify those who are “fit to study medicine,” as described Daniel Coit Gilman in 1878. The results of the work by Marcus and Tutton seem to suggest that selecting students based on their success in the premedical sciences may select against students who excel in empathy and humanism, characteristics that are equally important as, if not more important than, cognitive ability in the sciences.

One’s overall level of ability does not manifest itself in a single way and cannot be measured along a single axis. Lubinski suggests that there are at least three dimensions to ability: quantitative ability, spatial ability, and verbal ability. Within individuals, different combinations of these three dimensions will be manifested as different proficiencies and proclivities. An individual who excels in one dimension
of ability is not more intelligent than someone who excels in another. The concept of general intelligence is reflected in the combination of the three dimensions. Lubinski describes how these various abilities may affect educational outcomes: “People do not select educational tracks and occupations randomly. They do so, at least in part, on the basis of stable features of their personality, which include their specific abilities.”

A highly intelligent individual who scores higher in quantitative ability may likely be more comfortable in areas such as math and science; one who scores higher in verbal ability may instead be drawn to social science or the humanities. Conversely, a more verbally oriented person who, for whatever reason, begins his or her education in a scientific field may be more likely to shift out of that field to one that is more compatible with his or her inherent intellectual strengths. A series of research studies, using a unique tool to measure individual differences, confirmed this tendency to shift one’s area of study in the face of a mismatch of cognitive style.

Over a period of more than thirty years, Herman Witkin, a psychological researcher working with the Educational Testing Service, studied a phenomenon he referred to as “field dependence” and its association with educational preferences and success. To measure field dependence, a subject sits before a computer screen in a darkened room. On the screen is a quadrilateral figure that looks like a rectangle placed vertically in space but is actually tilted very slightly off vertical. Within the figure is a line that is clearly tilted off vertical. Using a control stick that rotates the line on its center, the subject is asked to reposition the line so as to place it in the vertical position. The question under study is, does the subject ignore the figure and place the line truly vertical, or does the subject use the figure as a field of reference, positioning the line parallel to the figure rather than truly vertical? By repeating the test with a more highly skewed figure, the researchers determine how far off vertical they can position the quadrilateral before the subject ignores it and positions the line independently in space. Those who follow the figure in positioning the line are “field-dependent,” while those who tend to ignore the figure are “field-independent.”

In their research, Witkin and colleagues followed more than 1,500 entering college students, administering the test and assessing their field dependence/independence at the time they entered college. They then followed these students through the four years of college and into either their work or their graduate education. They arrived at three principal conclusions:

1. Field-independent students tended to select an undergraduate major in the natural sciences or math.
2. Field-dependent students tended to major in education or related social sciences.

3. Either field-dependent students who initially selected a science/math major or field-independent students who initially selected an education/social science major were significantly more likely to change majors than students whose initial selection of a major matched their field dependency score.58

There has been debate among psychologists as to whether field dependence actually reflects a form of cognitive ability rather than simple visual tendencies unrelated to measures of intelligence.59 However, some of Witkin’s research results apply specifically to premedical students. Among Witkins research subjects were a substantial number of students who declared an interest in premedical studies early in their college career. (Since Witkin conducted his research at a time when fewer than 10 percent of medical students were female, he only reported results for males.) Consistent with conclusion (3) above, premedical students who were field-independent were significantly more likely to remain in premedical studies and subsequently to apply to medical school than were field-dependent students.

From other research he had done, Witkin was able to compare the personality characteristics of people who were field-dependent or -independent. The difference in personality profiles holds particular relevance to our discussion of the strengths and abilities we look for in physicians. As summarized by Witkin,

Field-dependent people are more attentive to social cues than are field-independent people. Field-dependent people have an interpersonal orientation: They show strong interest in others, prefer to be physically close to people, are emotionally open, and gravitate towards social situations. Field independent people have an impersonal orientation: They are not very interested in others, show both physical and psychological distancing from people, and prefer nonsocial situations. Finally, field-dependent and field-independent people are different in an array of characteristics that make it likely that field-dependent people will get along better with others.60

While acknowledging the debate over Witkin’s work, it is hard not to see its relevance to premedical education and to the process by which we select those students best suited to become physicians. Tutton’s comments cited above, “that students with high achievement in many components of the [premedical] curriculum tend to have personality profiles that seem inappropriate for their chosen careers as physicians,”61 appear also to apply to the process by which field-independent students tend to stay in premedical education and enter medicine, while their field-
dependent classmates tend to drop out of premedical education and enter other fields.

Whether it is the ability to be emotionally open and comfortable in social situations or the ability to empathize and show compassion, it should be clear that those with a higher level of ability in these areas will be perceived in a better light by their patients than those with less ability. Similarly, when they are evaluated by medical school faculty regarding their clinical skills, those who score higher on scales of empathy will also score higher on assessments of interpersonal skills when working with patients in a clinical context. This association was confirmed by two recent research studies. Hojat and colleagues from Jefferson Medical College administered what they referred to as the Jefferson Scale of Physician Empathy to 371 third-year medical students and then compared the students’ empathy scores with a global rating of their clinical competence provided by the faculty in their third-year clinical clerkships. The authors found a significant positive association between the two scores. Similarly, Stratton and colleagues administered a psychological assessment to 165 third-year medical students at the University of Kentucky and then compared these results with the students’ performance on a required comprehensive clinical performance examination involving a series of twelve encounters with standardized patients. As described by the authors, “Significant associations were found between the standardized patients’ rating of students’ communication abilities and their scores on one [emotional intelligence] and two empathy subscales.”

Consistent with our discussions above, professionalism for a physician implies both a certain level of cognitive ability in the sciences and a parallel, but distinct, ability in establishing a humanistic connection with patients. This brings up an obvious question: If medical students, selected principally on their proven cognitive abilities in the sciences, tend to be weaker in empathy, can they be taught how to feel and display empathy? This question has been addressed by a number of researchers.

Poole and Sanson-Fisher studied the effects of explicit empathy training on medical students in Australia. As described above, Australian students undergo a six-year medical curriculum, with the final three years completing the preclinical sciences and providing clinical education. The researchers administered formal empathy training to a group of 25 students selected randomly from a group of 45 entering the final three-year curriculum. They were able to identify clear improvement in empathetic ability immediately following the training: “Prior to training they had been hesitant, frequently avoiding emotional topics and tending to dominate the interaction; this changed following training. . . students permitted patients to do most of the talking and encouraged emotional expression; their responses generally indi-
icated understanding and concern for patients.” However, by the end of their clinical training, that empathetic behavior in the study subjects had decreased significantly and was only marginally better than the control group, leading the authors to conclude that “the educational experience during the clinical years appears to negate to some extent the earlier training.”

One has to question whether the students in the study by Poole and Sanson-Fisher were actually being taught to feel empathy, rather than being taught to act in an empathetic manner without actually personally experiencing empathy. Rees and Knight suggest that, in evaluating professionalism among physicians, we distinguish between attitudes and behaviors. One who acts compassionately does not necessarily feel compassion.

Whether patients can distinguish between compassionate or empathetic behavior that does not have an actual underpinning in true compassion or empathy and behavior that stems from innate empathy remains an open question. Benbassat and Baumal suggest that there is value in teaching students to act empathetically, even if they do not experience true empathy: “We believe that the ability to encourage a patient to convey his distress is a teachable skill, while the subsequent steps are mainly related to the personality traits of each individual student.” They emphasize the importance of teaching patient-centered interviewing skills and strengthening the student-physician/patient relationship by restructuring clinical teaching to permit the development of longer-term relationships between the student and the patient. However, the authors discourage the assessment of empathy as part of the admissions process because, “Even if admissions committees could identify an ability to empathize among entry-level medical students, the hospital environment would most likely eradicate this ability.”

In their analyses, these authors acknowledge the previous research demonstrating that medical students’ level of empathy as measured on standardized psychological instruments declines throughout the process of medical education. A recent paper by Newton and colleagues confirms this finding. They distinguish between “vicarious empathy,” in which a student has a genuinely visceral response to a patient or a situation, and “role-playing empathy,” in which a student is taught to act in an empathetic manner without experiencing the true sensation. Using the Balanced Emotional Empathy Scale instrument, they tracked 419 entering medical students at the University of Arkansas for Medical Sciences, spread across four entering classes. They administered the Empathy Scale at the beginning of each of the four years of medical school and found consistent declines over time. The decline was particularly steep following the first year and the third year of school. Interestingly, while all students showed nearly parallel declines in empathy over time, those
students choosing to go into the more patient-oriented fields of internal medicine, family medicine, obstetrics-gynecology, pediatrics, and psychiatry both entered school with higher levels of empathy and maintained higher levels over time than their classmates who chose more technically oriented specialties.\footnote{67}

Howard Spiro of the Yale University School of Medicine published an insightful commentary in 1992 about the role of empathy in medical education and medical practice. His comments were not particularly complementary toward the dominant model of medical education: “As I know them, college students start out with much empathy and genuine love—a real desire to help other people. In medical school, however, they learn to mask their feelings, or worse, to deny them. They learn detachment and equanimity. The increased emphasis on molecular biology to the exclusion of the humanities encourages students to focus not on patients, but on diseases.”\footnote{68}

Writing in 1983, Gregory Pence, a professor of philosophy and medical ethicist and a faculty member at the University of Alabama Birmingham School of Medicine, emphasized the importance of compassion in medical practice. Referring to the work of Aristotle, he stressed the importance for the physician of developing a sense of intimacy with a patient, intimacy that is “built on related moral qualities between listener and sufferer of trust, honesty, and the time and willingness to listen.” Pence shares the skepticism voiced above by Spiro:

The view [among medical educators] seems to be that the primary goal of medical training is to produce scientifically competent physicians and, as for compassion, well, it will be picked up by “osmosis” (perhaps “perfusion” would be more exact). . . .

Compassion in undergraduates is notoriously difficult to discover or measure, especially in the brief, episodic encounters of mass education between professor and student where future requests for recommendations may lurk in a student’s mind. But even if compassion could be accurately identified in undergraduates, the crucial problem remains of the great power of medical education to eradicate compassion.\footnote{69}

Coulehan and Williams argue that certain students enter medical school with a “natural immunity” to the tacit, implicit forces within medical education that lead to a loss of empathy among medical students over time. Despite the dehumanizing atmosphere of medical school, these “immunized” students “progress through medical school and postgraduate training while maintaining, even nourishing, an altruistic professional persona.” The authors suggest that, if it were possible to measure this natural immunity as part of the admissions process, we could avoid the current
skewing that favors “individuals who might turn into good scientists or technicians, but who have two strikes against them when it comes to becoming compassionate physicians.”

Psychologists will continue to debate whether qualities such as compassion and empathy are innate or can be learned. They will continue to differentiate between acting in an empathetic manner and genuinely feeling the quality of empathy. We can be sure, though, that psychologists, and for that matter philosophers as well, will generally agree that empathy forms a crucial underpinning for competence and professionalism on the part of physicians. The ability to feel empathy is not measured by the MCAT and is not reflected by grades in the premedical sciences—that is, unless the research of Tutton and Witkin, described above, is accurate and the ability to succeed in the premedical sciences is inversely related to one’s natural empathic abilities and sensitivities. If it were the case that, by selecting students for medical school based on the paradigm of premedical education established more than a hundred years ago, we are selecting against those students who are strongest in empathy and for those students who are weakest, we would be making a fundamental error. Quoting again from Kupfer’s research from 1978, “One interpretation of these findings would be that those ideal attributes found in the ‘good physician’ are derived from more than one dimension. . . . Therefore, identification of students who have both high MCAT scores and high empathy scores might represent one approach in predicting which students will make the best clinicians.”

The Need to Develop a Multidimensional Model of Medical School Admissions

We began this chapter with two quotations made nearly forty years apart yet concurring on the importance of both cognitive and noncognitive qualities in medical education and medical practice. By tracing more than forty years of research that followed the more recent of the quotations, we have found substantial support for this concept. Whether it is the “good physician” referred to by Kupfer or the ideal clinician described by Tutton, he or she will have at least two characteristics: an adequate knowledge of the scientific principles on which medical practice is based and the personal qualities such as character, empathy, and compassion that are requisite for optimal competence as a clinician.

In the MCAT we have a reliable and well-validated means of measuring the scientific knowledge gained from premedical education. As described in the previous chapter, performance on the biological sciences and to a lesser extent on the physical sciences component of the MCAT is a reasonably accurate predictor of future
performance in the preclinical sciences of medical school. These measures of science aptitude, however, tell us little about a student’s future performance as a clinician.

The admission interview, if done well, gives us some guidance about an applicant’s noncognitive strengths and weaknesses but is unable to provide an accurate prediction of a student’s future ability to communicate with patients and to gain their trust through the intimacy afforded by empathy and compassion. Clearly, we need ongoing research on how best to measure these noncognitive aspects of an applicant’s qualifications to be selected to enter medical school. However, for the purpose of discussion, let us assume that we have developed a measure of an applicant’s personality that is equally reliable and equally valid as the MCAT is in assessing scientific aptitude. How would we use these two measures of an applicant’s qualifications? The research cited above suggests that they need to be approached as dimensions that are largely independent. Using these two measures, we could therefore place applicants somewhere in a two-dimensional grid of qualification such as that depicted in figure 5.1.

In figure 5.1, the vertical axis represents a continuous scale of scientific aptitude, evaluated by measures such as the MCAT score or premedical science grades. The
horizontal axis represents a continuous scale of personality strengths, assuming we could eventually develop such a measure. The area within the square represents the pool of applicants to medical school. Within this pool there will be some people who simply are not qualified, due to gross inadequacies in their science preparation. These are represented by the diagonally shaded box along the horizontal axis. No matter how great the strength of their personality, these applicants are likely to be unprepared to undertake the study of medicine. While some of these students might eventually succeed in the first two years of medical school, the probability is high that many of them would fail one or more of their preclinical science courses. It was this population that the MCAT was first developed to identify, as described in the previous chapter.

Similarly, there is a somewhat narrower shaded box along the vertical axis, representing those applicants with shortcomings in one or more aspects of their personality that disqualify them for medical practice. While some of these students could possibly complete medical school and become medical researchers with little patient contact, their personality weaknesses make them inappropriate for working with patients. Through the admissions interview and information provided by premedical advisors, these students also need to be identified.

In the upper right corner of figure 5.1 is a smaller, lightly shaded box, representing those medical school applicants who score the highest on both scientific aptitude and strength of personality. Clearly these are the applicants medical schools want the most. They have the firmest grounding in science, and we can reliably predict that they will be seen by their patients as representing the highest level of medical professionalism. Of course, the number of applicants represented by this small box is inadequate to fill the entering class of all the medical schools in the country. How should we expand the box of qualified applicants to a size sufficient to fill the entering class? There are two ways to do this: the way we have been doing it for nearly a hundred years, represented by figure 5.2, and an alternative model, represented by figure 5.3 and based on the realization that we need to evaluate the fitness of an applicant to study medicine using at least two equally important dimensions of quality.

Since the 1920s, when for the first time the number of applicants to medical schools exceeded the number of available places in the first-year class, we have used scientific aptitude as a nearly continuous measure of an applicant’s fitness to study medicine. Premedical performance in chemistry, biology, and physics is the principal measure admissions committees have used to rank-order applicants. Extracurricular activities, comments of the premedical advisor, and results of the admission interview provided additional information at the margins. Despite a widespread
recognition that character traits such as empathy and compassion play a major role in medical professionalism and the quality of care, the absence of a reliable means of assessing these characteristics left admissions committees little choice but to select those students who were the least likely to fail the preclinical sciences classes encountered in the first two years of medical school—the traditional measure of fitness to study medicine.

The research described in this chapter, however, suggests a fundamental weakness in this approach. Fitness to study medicine is determined by a combination of an applicant’s scientific aptitude and his or her personality strengths. Only by including both dimensions in the evaluation process can we get the most complete picture of an applicant’s qualifications.

Figure 5.2 provides an alternative approach to expanding our concept of “fit to study medicine.” Beginning in the upper-right corner of the diagram, it identifies the set of most highly qualified applicants, but not as a horizontal rectangle using scientific aptitude as means of ranking applicants. Instead, it expands the area of the pool by maintaining the shape of an isosceles triangle, moving down each axis equally and thereby giving equal weight to an applicant’s scientific aptitude and his
or her personality strengths. Under this alternative model, there are some applicants who would be admitted to medical school who would not have been admitted under the traditional model—those with relatively lower scientific aptitude but relatively greater personality strength. Conversely, there are some applicants who would have been admitted under the traditional model but who would be denied admission under this model—those applicants who are relatively higher in scientific aptitude but with relatively weaker personality strength.

Which model would produce better physicians? Could we trust as our physician someone who got a B in organic chemistry in college, when we might instead have a physician who got an A? Would a 10 on the biological sciences portion of the MCAT mean that a physician is not as qualified as his or her colleague who got a 12?

Research suggests that patients will tend to put aside these questions and ask instead, “Can I trust my physician? Can she or he really understand what it feels like to be sick or injured?” For most patients, that is what makes a good physician.

No matter how much we agree that personality and strength of character are important in selecting students for medical school, we still don’t have a means of assessing these noncognitive qualities in a way that is both reliable and feasible. How-

*Figure 5.3. An alternative approach to evaluating qualifications for admission to medical school*
ever, we can address the common concern that de-emphasizing academic performance in the premedical sciences will weaken the clinical and professional capabilities of the physicians we train. By looking at research reporting the outcomes of programs that have placed less emphasis in the admission process on attainment in the premedical sciences, we can determine what effect such nontraditional approaches have had. These programs are of two general types: those specifically intended to increase the ethnic diversity of the medical profession, and those that have adopted a nontraditional approach to medical education without ethnic diversity as a singular goal. I discuss these programs in the following chapters.