CHAPTER 1

COLLEGE ADMISSIONS AND TESTING

Waiting list. That was the long-awaited outcome of my application to Yale. It was better than the outcome at Harvard, which was a rejection, and not as good as the outcome at Princeton, which was an acceptance. But I wanted to go to Yale.

I was to have one opportunity that few applicants would ever get. I was eventually admitted to Yale off the waiting list, and went there, later to graduate summa cum laude, Phi Beta Kappa. The opportunity I had was to find out why I was wait-listed. Right after graduation, I was working as a special assistant to the dean of admissions at Yale. With the encouragement of some of my colleagues, I sneaked up into the dusty old attic where old admissions records were kept. I did not read my whole admissions file—I felt too guilty—but I did not feel guilty enough to pass up reading my Yale interview report.

I remembered well my interview visit to Yale. It was a disaster. It was pouring cats and dogs. I wanted to major in psychology and so visited the psychology department. They hooked me up with a postdoctoral fellow who was implanting electrodes in cats’ brains. I was revolted. They had just painted the psychology department, and I got bright blue paint on my raincoat that never did come off. And my interview was worst of all. In the in-
terview report, which I discovered up in the attic, the interviewer said I had a “flakey personality.” And did Yale want a flakey personality? When I finally did get in, it was through the intervention of the admissions officer for my area, Bill Robinson, who saw something in me.

Great schools do not always produce flawless leaders. Harvard Business School produced Jeffrey Skilling, former CEO of Enron and convicted felon. Yale College, my own alma mater, produced George W. Bush, whose invasion of Iraq was based on intelligence that even he later admitted was faulty. Yale Law School produced Bill Clinton, whose inability to control his self-destructive tendencies resulted in an impeachment that severely undermined the legacy of his presidency. Dennis Kozlowski, who went to college at Seton Hall, a few miles away from where I was going to high school at the same time in New Jersey, was convicted of looting his company. When he was CEO, Tyco paid $1 million for his wife’s fortieth birthday party. The party, billed as a shareholder meeting, featured a statue of Michelangelo’s David urinating Stolichnaya vodka. Kozlowski also had Tyco pay for his $6,000 shower curtains.¹

Such examples of flawed behavior by very successful college applicants, however, only imperfectly capture the problem of college admissions. At least as problematic are the many people who look as though they have great potential at age seventeen or eighteen but don’t look so great twenty years later, along with the potentially great ones who get away. In talking to a high-level executive at a major investment bank, I mentioned our desire to enhance admissions at Tufts University. His response, based on his twenty-five years of experience on Wall Street, was that tests like the SAT and the ACT, as well as college grades, predicted quite well who would be good analysts. That is, they predicted the technical skills needed to evaluate various investments. What they did not predict as well was who would be able to take the next step—who would have the capacity to envision where vari-
ous markets are going, to see larger trends, and to make decisions that went beyond individual stock or bond picks. Are great schools missing something in how they admit students and how they teach them?

All great schools seek students who will enhance the excellence and diversity of their classes. They want students who will shine in the classroom and then excel in later life, making the world a better place in which to live. Are we really doing the best we can, as a society, in selecting and developing the best generation of leaders? Do the current elements of a college application identify these traits and abilities in a maximally effective way? Or is there something lacking in what we are doing? That is, are we always going to be stuck with a few bad apples, no matter our methods, or is there something wrong with the way we pick and sort the apples in the field?

I believe that the problem is not just with a few bad apples, or with our sorting methods, but with society: at some level, we all contribute to it. Some parents push their kids relentlessly to get better scores on standardized tests, often paying obscene fees to tutors ($500 and more a session in some New York suburbs). College and university administrators contract with testing companies to use these standardized tests so that they can admit applicants they believe to be qualified—as well as, in some cases, to compete in U.S. News and other ratings sweepstakes. College and university faculty want students with high test scores who will do well on the professors’ own tests. Students dutifully take the tests, and they or their parents pay for the tests that may get them into, or keep them out of, the colleges they dream of attending. The test companies continue to market roughly the same tests they have been selling for a century. And the government is blind to what seems to some to be the monopolistic hold that two test companies have on the college admissions market.

Here, as a teaser, are the six main claims of this book:
- Tests such as the SAT and the ACT are not "bad," but rather incomplete. They measure memory and analytical skills, which make up only a fraction of the skills that are important for college and life success.
- The solution to the incompleteness of such tests is not to replace them, but to supplement them or to have tests that measure not only what the traditional tests measure, but also other qualities.
- Other skills that are important to measure are creative skills, practical skills, and wisdom-based skills. To succeed in school and in life, one needs creative skills to generate new ideas, analytical skills to ascertain whether they are good ideas, practical skills to execute the ideas and to persuade others of their worth, and wisdom-based skills to ensure that the ideas help attain a common good, not just selfish gain.
- Using tests of these additional skills increases prediction of both academic attainment and meaningful participation in extracurricular and leadership activities in college.
- Using such tests simultaneously with standardized testing reduces, potentially substantially, ethnic-group differences in overall test performance.
- In the end, such enhanced admissions procedures benefit candidates for admission, the colleges to which they apply, and society at large.

Some, such as Frank Schmidt, John Hunter, and others, have suggested that we keep using conventional tests because they worked in the past and they work in the present.² I agree. They do work, at some level. They predict modestly to moderately many different kinds of skills, from getting A's in school to following complex regimens with prescription drugs. They also can help admissions officers sort out students who may not be able to do the academic work in a given college from those who are more likely to be able to do that work. The problem is that today,
as in the past, their prediction is far from perfect. Remember, early medicine and early telephones worked too. They just did not work all that well. And the reason they developed rapidly and standardized testing has not is simple—competition.

When the telephone companies were deregulated and AT&T lost its monopoly, telephone technology and service began to expand at a fast clip. In the computer-related fields, competition meant that businesses like Control Data Corporation (CDC) that did not innovate died, while those businesses that did innovate, such as International Business Machines (IBM), survived and even thrived. Drug companies today compete for ever greater shares of markets. But two companies largely control the market for college admissions tests, with one substantially larger than the other and with each having areas of control rather tightly defined by region of the country. These companies would themselves benefit from increased competition, as does any company that wants to produce the best possible consumer products.

IMPROVING THE SYSTEM: VALUING ADDITIONAL SKILLS

Although the current system of admissions is good, it can be substantially better. We can do a much better job of college admissions, as well as instruction and assessment, if we think about student abilities in a broader way than we have—in particular, by valuing, assessing, and teaching for analytical, creative, practical, and wisdom-based skills as well as for memory-related ones. In particular, I propose that students should be admitted to college on the basis of their potential for future leadership and active citizenship, at whatever level of society (from the family, to the workplace and other communities, to the world), by taking into account, among other criteria, their having the academic knowledge and skills necessary for success in college work.

I define leadership here not in the sense of achieving a level of authority, but rather as making a positive, meaningful, and hopefully enduring difference to the world at some level. Many
current admissions processes seek an understanding of a broad range of talents, but test scores and school grades, because they are quantified, may get proportionally more weight than they deserve in a holistic admissions process. Such tests are somewhat narrow in the spectrum of knowledge and skills they measure, and need supplementation by assessments that draw on broader theories of human abilities and competencies. Such broader tests would help to create systems of admissions that are more nearly equitable for the individuals applying to college, including those from diverse socioeconomic backgrounds who have hidden talents to offer, and would improve outcomes—both in terms of academic quality and diversity—for universities and society as well.

This broader vision of the ideal college student should also influence the education of these students. Although colleges often treat admissions and instruction or assessment as distinct and autonomous operations, they need to be coordinated so that instruction matches the abilities selected for in admissions. For example, if a college seeks creative students through its admissions program, then it must also help instructors teach in ways that encourage and value creative thinking, lest students find themselves in an environment incompatible with the very skills for which they were admitted.

WHY IS CHANGE NEEDED?
There is something lacking in the way college admissions are conducted. We are not admitting people as we ideally should and could, and the academic disciplines are not doing a perfect job of cultivating students’ skills, including their ethical ones. It is not that we are doing a bad job; it is that we can do much better. Neither testing companies, nor admissions officers, nor any other group is entirely to blame. Rather our society, as a whole, has created a system of interlocking parts that do not work together as well as they could—and it will be hard to fix because so many people do not even recognize that there is a problem.
Kenneth Lay, who like Jeffrey Skilling was both a former CEO of Enron and a convicted felon, had earned a doctorate and was a professor of economics. Did academic knowledge save him from helping to destroy a company with thousands of employees and customers? Apparently not. Of course, there are many well-educated people who are wonderful citizens and great leaders. But to the extent they are either, it may be not because of their knowledge base, but rather because they understand what to do with it.

Until the 1960s, most students were admitted to selective colleges on the basis of their parents' social class. It was believed that one’s social class would predict in great measure one’s potential for future positive leadership.

In the 1960s, Inslee Clark, dean of admissions at Yale, and others involved in college admissions had a new vision—one of an elite chosen primarily on the basis of merit rather than largely on privilege. No longer would it be enough merely to come from an affluent family, or to have acquired an education at one of the elite private schools in the nation. Some great students might come from well-endowed families and elite private schools, but then, some might come from very poor families and public schools with few resources. So the colleges started giving test scores and grades more weight in their deliberations, and parental wealth and privilege less. The result was a new generation of leaders who were supposedly chosen on their merits rather than by the luck of the draw at birth.

Given the leaders we have—ones who led Lee Iacocca to write the book Where Have All the Leaders Gone?—we have to question whether something went wrong somewhere along the line. Actually, several things went wrong, and these problems are still with us today.

**We Need to Identify and Nurture Potential Leaders**

The first problem was that test scores correlated highly, although not perfectly, with social class. This correlation is not fixed;
rather, parents of children from upper socioeconomic classes inadvertently increase this correlation. They, unlike less well-off parents, can afford to spend substantial sums of money having their children tutored for the tests they will have to take, so that these students come to have an edge in the testing sweepstakes. It is understandable that well-off parents would pay for such tutoring: they are trying to do the best for their children. If only test scores well predicted future leadership, then we might be back to our hope that family legacy could predict who will make a positive difference to the world in the future. But they offer no magic numbers for predicting future leadership or anything else.

This problem is an example of what is sometimes known as a Matthew effect. The Matthew effect derives from a statement in the Bible: “For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath” (Matthew 25:29). In other words, the rich get richer, and the poor get poorer.

The late Robert Merton, a sociologist at Columbia University, applied this principle to scientists: those at well-known schools who already had a good reputation received more resources, whereas those at poorly known schools who had little reputation received less, resulting in the further derailment of their careers.\textsuperscript{5} But the Matthew principle applies as well to college admissions. If your parents can send you to strong schools and pay for tutelage, they can help bring you to a position where you are more competitive for college admissions; if your parents cannot pay, it becomes increasingly difficult to gain access on the basis of traditional measures of merit. Good admissions officers will take these factors into account and consider access to resources when evaluating applications.

Why should a correlation of test scores with socioeconomic status even be a problem? After all, if students from higher socioeconomic strata are able to perform better in college, then why not simply let them in, that is, why not simply follow the college
admissions tests, which would accurately predict their performance?

There is no simple answer to this question and the answers one provides are a matter of policy, not psychology. But I believe there are three reasons why the correlation is problematic. To understand these reasons fully, you will need to imagine a society that educates well those people with blue eyes but poorly educates people with brown eyes, having decided that eye color is a good measure of the educational and perhaps other resources to which a person is entitled. (In our society, of course, historically we have used gender and skin color, which are equally arbitrary.) When the children reach high school, society gives them a test. The blue-eyed children outscore the brown-eyed children by a large margin and then, because college admissions are largely determined by test scores, the blue-eyed children dominate the numbers in college.

The first reason I can think of for not settling for such an outcome is that if one makes decisions largely on test scores, as partial proxies for socioeconomic status (which itself may be a proxy for other things), then one reduces one's chances of creating a more equitable society. There will be those who will say that society has been fair, because it gave preference to those who deserved it. Unsurprisingly, many, but not all, who think that way will have blue eyes. Many of the brown-eyed adults may say that their children were never given a chance, but the brown-eyed adults will have less money and power in society, and will not be well heard. Moreover, some of those who do hear them will view their complaints as sour grapes. Without some kind of admissions program that looks for strengths in the brown-eyed children, the society will be stuck with a hereditary hierarchy that seems merit-based—at least to those with blue eyes, including educators, scientists, government officials, and others who are perfectly rational on many matters not pertaining to the continued social empowerment and well-being of themselves and their children.
The second reason is that the brown-eyed children may actually be better in some skill areas than the blue-eyed children. This superiority will probably not show up in assessments created by typical blue-eyed adults, who have come to view as valuable those skills in which they and their children excel. But without efforts to devise measures that may do justice to the full range of skills of all children, one may in effect create tests skewed to favor those with blue eyes. Moreover, the teachers, who are also primarily blue-eyed or at least trained by those with blue eyes, will value the same narrower range of skills. So they will construct assessments that match the college admissions test, in the same way that the college admissions test may well originally have been created to value what the blue-eyed teachers, or those teachers trained by those with blue eyes, value. The tests will seem fair because the system is created to value what it is that blue-eyed people, on average, are presumed to do well, but not what brown-eyed people, on average, are presumed to do well.

The third reason, already implied, is that the criterion may be as biased, in a sense, as the predictor. When people are hired for high-level jobs, employers often consider what college the student went to (for example, Harvard versus Stinky U), and what kinds of grades the student received. But schools and grades are a highly imperfect predictor of job performance. The recent debacle on Wall Street is an example. It was created largely by individuals with excellent grades at terrific business schools. It was their ascent into positions of great power that enabled them to bring down not only Wall Street, but much of the economic system of the United States and the world as well. Indeed, it is a testament to the system's inherent bias that, when some of the "blue eyes" were fired for extraordinarily incompetent performance, they received severance packages—and many still there are receiving bonuses—in excess of what many brown-eyed people will make over their entire lifetimes. Heads they win; tails, others lose.

It is important to note in this consideration that high-school grades have correlations with socioeconomic status that may be
as great as or even greater than those of the standardized tests. So ditching the tests in favor of considering only school grades might solve some problems, but not that of the correlation of academic performances with socioeconomic status. The correlation of socioeconomic status with high-school performance is not terribly surprising. Educated parents can help their high-school student children in ways that less educated parents cannot. In my own case, neither of my parents had graduated from high school. They could help me in school, but perhaps not in the same way as could the parents of some of my classmates, who held doctorate degrees.

We Need to Reward, and Teach, a Broader Range of Skills

The second problem is that the tests were narrow in their conception. In the early twentieth century, when the tests were first devised, perhaps the abstract, academic skills they measured served as a somewhat reasonable basis for distinguishing more able white males of privilege from less able ones. But by the early twenty-first century, the skill set one needs for success in college and life has substantially broadened. Academic knowledge alone will not get one through; the world simply changes too quickly. (More on this later.) For this reason, many colleges use holistic admissions practices that take into account credentials beyond just academic knowledge.

It is worth emphasizing that the system we have now was made to create equity, not to destroy it. In earlier times, the main way one got ahead was through family connections. Tests were designed to create an objective measure of potential that would move beyond family ties. The procedures may have made sense in the early twentieth century, when the only people taking the tests were the children of privilege, so one could distinguish better among them who had more potential. But as the range of students taking the tests has increased, and their backgrounds have become more diverse, what had worked in the early twentieth century simply no longer works as well. One can no longer
assume that almost all the test takers have similar upbringings, backgrounds, and opportunities for schooling. New tests are needed to reflect this new reality. But none of the new tests in development has yet been widely accepted.

The third problem is that the tests transformed secondary schooling, and not necessarily for the better. Students have come to spend more and more time preparing for tests, and less and less time learning lessons that may be meaningful to them in later life. Education has come, in some measure, to be replaced by gamesmanship. Music, art, physical education, and even, in some schools, social studies and science have gone by the wayside if they are not formally tested by the states.

**WHY TESTING HAS CHANGED SO LITTLE**

How has the testing game come to take over not only schooling, but even much of society? The route to the “takeover” has been a series of unfortunate collateral consequences of testing that has entrenched it ever more firmly in society’s psyche. The problems that have resulted have little to do with the way admissions officers think. Admissions officers are, for the most part, very skilled at what they do. Rather, the problem is a societal one, reflecting the way society in general reacts to decision-making practices regarding college.

**Traditional Tests Seem Precise**

Tests yield quantified, seemingly precise measures of students’ abilities. Consumers love tests’ apparent precision of measurement. And it is much easier to make a decision relying heavily on numbers than to make one relying heavily on seemingly subjective data, such as teachers’ letters of recommendation or lists of extracurricular activities—even if the scores on these tests are nowhere near as valid as they appear to be. Skilled admissions officers, therefore, take into account the subjective as well as the objective factors, recognizing the need for a holistic evaluation of
each applicant. They resist the temptation to do their job the easy way—by the numbers.

Anyone is susceptible to overinterpreting numbers. Some years ago I was giving a talk to an audience of people who drilled for oil. They were interested in finding people who would be skilled at predicting where oil could be found. As I was talking about our work, someone in the audience raised his hand and made a curious point. He said that their company had the same problem with clients that we had with admissions. Clients would often prefer that they drill for oil in places where there were quantitative indices indicating the likelihood of oil, even if those quantitative indices were known to be largely invalid. They would prefer to go with quantitative rather than superior qualitative information based on the judgments of experts, simply because it was associated with numbers. And they got a lot of dry holes.

Successful Test Takers Seem More Attractive as Applicants
People are attracted to others who are like themselves. We tend especially to like others who are similar to us in interpersonal attractiveness, who share our interests, who are in our ethnic group, and so forth. The tests that have been so highly valued since the 1960s have yielded a new generation of power-holders who look for others like themselves—that is, others who have high test scores.

During the era in which most people going to elite colleges were from socially elite families, admissions officers were themselves mostly from elite families. So they looked for people with similar backgrounds. Today, to be admitted to elite colleges, it helps to have high scores from either the SAT or ACT. As a result, some admissions officers had relatively high board scores, or work for faculty who did. So some of these admissions officers may look for applicants who, like themselves and their mentors, will continue the legacy they have created. Such methods of producing social stratification are hard to change. But they explain
why increasing the diversity of admissions staff can increase the
diversity of the matriculated student body.\textsuperscript{7}

Which system for selecting college students is better—test
scores or social class and wealth? It depends on whom you ask.
People in positions of power tend to value whatever attributes
got them into those positions. If you ask a Harvard professor
who is making $150,000 a year and who has published 150 arti-
cles in her career, the professor might wax enthusiastically about
the advantages of the standardized tests that helped her get to
where she is. If you instead ask an entrepreneur who dropped out
of school and is making $150 million a year, that individual
might laugh in your face. She might well think that money is a
better basis for making decisions than test scores.

Someone else might question both test scores and wealth as
bases for decision making. Mother Teresa and Mahatma Gandhi
led lives of relative poverty but made an enormous positive dif-
fERENCE to the world. They or those who admire them might ar-
gue that much more important than test scores or family wealth
is the set of positive ethical principles one brings to one’s life, and
one’s willingness to act on them.

In fact, there is no one perfect set of criteria for admissions—
family wealth or status, test scores, ethics, all might form bases
for people’s theories of what one should look for in a qualified
college applicant. It is important, however, to consider some his-
torical background on the issue. And in the past, one of the most
important criteria was gender.

When I started at Yale, it was an all-men’s school. The first
year that women were admitted, there was widespread outrage
among alumni. Yet within a short time, women were, on aver-
age, outperforming men in their academic work. And when I
became president of the American Psychological Association, I
would go to meetings of the board of directors and look at the
pictures of all past presidents of the association. It was pretty ob-
vious that they were overwhelmingly men. Some might take this
to indicate that men were more fit than women for such an es-
teemed position. But given that graduate schools for most of the history of the country did not admit women, what is truly surprising is that any women made it to such a position.

Other groups have experienced discrimination in the United States. To take just one example, if you were born in 1800 as a slave, and had an IQ of 160, it bought you little: you still died a slave. By contrast, if you were born the first son of a plantation owner, and had an IQ of 80, it mattered little: you still inherited the plantation. The law of "primogeniture" guaranteed that the inheritance passed to the eldest son. And no doubt the eldest sons felt that the law was remarkably fair in its recognition of their special role in their family and society.

Research by Claude Steele shows that socially defined race still matters in our assessments of ability, independent of the actual tests. Steele and his colleague Joshua Aronson have done a series of studies in which both white students and black students were asked to take a difficult test of verbal ability. When the participants were reminded, even subtly, of their socially defined race as part of the test-taking process, there was a greater difference in performance between whites and blacks than when they were not so reminded.8

It may be hard for some of us to believe that variables such as test scores, parental wealth, sex, socially defined race, and caste still affect one's possibilities in life. But imagine for a moment that the biggest determinant of an applicant's place in society in general, and in the college sweepstakes in particular, was, say, height. What then?

This thought experiment is less absurd than it might seem. First of all, our society does use height to determine social outcomes. Taller people are, on average, more successful than shorter ones, particularly among males. CEOs, army generals, and other successful people are more likely to be taller than average height, perhaps because people seem to respond better to authority in tall people. Second, height actually has an advantage over test scores: everyone knows exactly what it is. For example,
I am about five feet, eleven inches tall. You know what that means, right? But do you really know what an IQ of 125 means, or an SAT score of 580? And are they really different, or do IQs and SATs largely measure the same thing? (Research suggests that they are indeed closely related measurements.)

Some years ago, my colleague Douglas Detterman and I edited a book in which two dozen experts in the field of intelligence were asked to define intelligence. How many different answers did we get? You guessed it: two dozen. Even experts cannot agree on what intelligence is. But there is virtually no disagreement regarding the nature of height.

Second, height is the same regardless of the instrument with which it is measured. You could measure me with any of several tape measures, yardsticks, rulers, or whatever, and within a tiny margin of error, I would be five feet, eleven inches. If you measured me with a metric tape measure, it would come out at about 1.80 meters, but that still converts perfectly to five feet, eleven inches, with no loss of information. Ability tests have more complicated outcomes. You can give me half a dozen different ability tests, and I most likely will end up with half a dozen different ability scores, some of which may not be very close to each other at all.

Third, height has a reassuring consistency over time that is lacking in ability-test scores. I am the same height today as I will be until I shrink in my golden years. In contrast, one can take an ability test one day, and the same ability test a day, a week, or a month later, and there is much less guarantee of consistency. On tests like the SAT, variations of more than a hundred points between one testing to another are, while not common, far from rare, and fifty-point variations are quite common.

Fourth, some students cheat on standardized tests. They write their answers on their hands or inside their shirtsleeves, store them in the bathroom and then seek permission to go there, bring in hidden PDAs, or whatever. With height, all of these temptations to cheat are reduced. Sure, students can cheat by
wearing platform or elevator shoes, but such devices are easy to detect.

Fifth, you can't buy height. Conventional test scores are susceptible to modification if parents can afford to buy their students books, courses, or just a sound education. With height, you've got what you've got unless you have the misfortune to have been severely and chronically malnourished. In that case, college probably won't be in the cards anyway, at least in the United States.

So there you have any number of reasons why it makes sense to use height for admissions to college. Suppose that we do in fact adopt the simple, easy-to-use height test. Now, to get into Harvard, you might have to be about six feet, eleven inches. To get into Yale, perhaps you would need to be only six feet, ten inches, and one goes down all the way to Squeedunk, for which admission requires one only to be three feet, one inch, and Podunk, which requires a height of a mere three feet even. What would the result be?

Almost without doubt, twenty-five years later most of the CEOs, army generals, and other people of high position would be tall, and most of the flunkeys of the society would be short. And what would this prove? That society favors tall people, in much the same way that it favors rich people over poor ones, men over women, whites over blacks (or, in some societies, blacks over whites), people of one religion over people of another, higher castes over lower castes, and so forth. That is to say, it would prove nothing at all other than that people can devise their own methods of social stratification, as well as the means to enforce the self-fulfilling prophecies that they have created.

**Ignoring Test Scores at One's Peril**

In some schools, admissions officers are reluctant to admit students who do not have high test scores. If the students are unsuccessful, the admissions officers are afraid that they, and not the students, will be blamed.
Indeed, every year some students in any given freshman class are identified as being at risk for flunking out, and at some schools, deans may look back to their admissions records for clues as to what went wrong. Why are these students failing, whereas others are not? Suppose those examining the admissions records discovered that the students who were flunking out were all admitted because one particular admissions officer had given them high ratings, whereas no one else had. That admissions officer would justifiably fear for his or her job. So people in positions of authority, whether in a college, graduate school, law firm, or anywhere else, worry about the decisions they make, because these decisions matter. What can they do to protect their reputations in case the people they recommend end up failing?

One thing they can do is use decision-making criteria that will help insulate them from criticism. Test scores are such criteria. If students flunk out or fail to adjust, the admissions officers or other decision-making personnel can blame the tests for leading them astray. But if they ignore the test scores, they may feel that others will place the blame squarely on their shoulders. And because test scores do correlate positively with academic success, those who make admissions or hiring decisions actually are taking a risk when they admit or hire people with lower test scores. Afraid of their own perceived future culpability, they may decide that the potential costs of taking such a risk are greater than the potential benefits.

**Pressure from Published Ratings**

It would be hard to overestimate the effects that published ratings have on the behavior not only of colleges, but also of elementary and secondary schools. Many, and probably most, elementary and secondary schools are shamelessly teaching to the standardized tests their students will have to take. They have to—the No Child Left Behind Act has given them little choice. The federal and state governments have placed a lot of pressure on schools to show that students have mastered a wealth of aca-
demic knowledge that will not necessarily matter greatly in their later lives. Principals risk losing their jobs and school boards may well lose control over their schools if their students under-perform.

In addition, because the results of various districts on these tests are publicized, they become excellent predictors of real-estate values, which are driven up in districts where test scores are higher. I know this from firsthand experience. Recently, my wife and I were looking for a house in the Boston area. Test scores in a given community were a terrific proxy for real-estate prices. If you knew the test scores, you pretty much knew the house prices, and vice versa.

While tests potentially divert schools from their responsibility of fully educating students, *U.S. News and World Report* and other magazines that publish ratings of colleges and graduate schools encourage the schools to set goals that raise ratings rather than educate students. Test scores matter a lot for these ratings, and moreover, they are more manageable than other criteria that are used to evaluate the schools, such as alleged national reputation, whatever that is. So schools work hard to increase their mean test scores, thereby giving more authority and power to the publishers of the tests. It is a very good time to be a test publisher, because if you can get your product entrenched in the marketplace, it is likely to produce a lot of money, regardless of what it happens to measure.

One could argue that tests are a means toward accountability, and that all the magazines and newspapers are doing is making these indices of accountability public. This argument is a good one, assuming that the measures of accountability are also good ones. And presumably, the measures are not terrible. They do not require, for example, knowledge of techniques of black magic or voodoo. But, as I will argue later in the book, the tests we use tend to be incomplete, and they often measure types of knowledge that will not necessarily be of great use in the workplace, or even in life more generally.
Colleges are eager to improve their U.S. News and World Report ratings as well as ratings by other media. Higher test scores improve ratings, whereas lower test scores risk lowering ratings, and, possibly, risk the jobs of those responsible for the lower ratings.

Superstitions and Self-Fulfilling Prophecies

Every once in a while, I am invited to give a talk in a highly desirable location, perhaps Arizona or New Mexico. The problem is that, after I give my talk, they send me home and rarely invite me again. I keep thinking it would be nice to get another invitation so I could spend more time in one of the more desirable vacation spots.

Suppose I open up a service that guarantees rain. Given the increasing scarcity of water resources, such a service is likely to have a bright future. I promise municipalities, states, really, anyone who will invite me, that if they contract with me to make it rain, I will make it rain or I will give them double their money back. So I finally wangle out that return invitation to visit Arizona.

When I get to Arizona, instead of giving a talk on college admissions or leadership or something about which I really have some knowledge, I do a rain dance. The question is whether, later in the day, it will rain. Given that I am in Arizona, that is scarcely likely. So the town that has invited me asks for double its money back. I explain that there has been a grave misunderstanding: this is Arizona; they can scarcely expect the rain dance to work in just one day. In an arid environment like this one, it may take weeks or even months for the rain finally to come. So every day I do the rain dance in the morning, and sightsee during the rest of the day. Eventually, it rains. I congratulate the people of the town on their good judgment in inviting me to make it rain, and I pack my bags and leave.

This little story may sound silly, but it is exactly how superstitions work. If you do a rain dance long enough, eventually it will
rain. And if it never does, you won’t be alive to tell the story much longer. You may reply that you don’t believe in rain dances. I don’t either. But the chances are you have some superstition that is equally powerful. For example, approach the first-floor elevator in an apartment complex or a busy office building and wait by the elevator. Sooner or later, someone in a hurry will come and push the elevator button, even though it is already lit, indicating that someone has already summoned the elevator. Why would anyone push the button when it already has been pushed? One reason is that elevators provide what psychologists call a 100 percent reinforcement schedule. If you press the button, the elevator always comes. So even though pressing an already lit button has no discernible effect, people keep pressing because they are always rewarded for doing so.

I may sound as though I am making fun of the superstitions of others, but I have my own. I wear a medal around my neck that my parents gave me forty-seven years ago. They told me it would bring me good luck. Does it? I have no idea. But I keep wearing it because I feel that, in general, I have had a good life, and the cost of wearing it is very small. Probably taking it off would have no effect, but what if it did? And that is precisely how superstitions maintain themselves, including in the academic realm.

Once society comes to believe that high test scores are a necessary condition for student success, its members continue to hold this belief even in the absence of evidence. When I was at Yale, some of my colleagues used to tell me that students with scores below 650 would not succeed in our graduate program. They pointed out that, in fact, all successful students had scores above that level. They were right: at the time, we refused to admit students with lower scores, so we never found out how they might perform. In colleges and universities, we have our own superstitions, and over time, we come to believe them, often thinking they are rational rather than fanciful.

Some superstitions are benign. For example, no one is hurt by my wearing a medal around my neck. But other superstitions
have serious consequences. As a result of our overuse of testing, many students never get a chance to show what they could do if only they were given the opportunity. So students who do not fit the particular mold that the tests create suffer as a result of our certainty that the test results mean much more than they do.

I have experienced the effects of such superstitions personally. When I was young, I did poorly on IQ tests. I would like to believe that this was because of test anxiety, but who can say for sure? In the 1950s, when I was growing up, the elementary school I attended gave group IQ tests every couple of years. As a result of my low scores, my teachers thought I was stupid and I did too. They never came out and told us our IQ scores, but one could tell from the way teachers acted. In first grade, I was a mediocre student, which made my teachers happy because they got what they expected. I in turn was happy that they were happy, and in the end, everyone was quite happy. By second grade, I was slightly worse as a student, and in third grade, still worse. This is a fairly typical pattern. It is sometimes referred to as “cumulative deficit.” Once low expectations set in, every year one performs a bit worse than the year before. Eventually, one is labeled a perennial loser.

Was I really a perennial loser? To the extent that there was a superstition that the low IQ scores ensured my poor academic performance, the superstition created a self-fulfilling prophecy, of which I was the victim. So academic superstitions are not victimless. They affect those about whom the superstitions are held. They also affect schools and society, because they cause students to underperform, and for talents to be lost to society.

I was lucky. In fourth grade, I had a teacher, Virginia Alexa, who saw beyond the tests in the same way that some college admissions officers do. She had high expectations for me, and she conveyed these high expectations to me. Just as I wanted to please my first-, second-, and third-grade teachers, I wanted to please her. In fact, I was extremely taken with her, and remember regretting only that she was so much older than I—and married.
So in fourth grade, I became an A student. My entire future trajectory changed as a result of just one teacher. But I have often asked myself: What would I have done if she had not been my teacher? Where would I be today? And when I was working in college admissions, I wondered how many of the students applying might have had far stronger records, if only they had had, as I did, a teacher who believed in them.

If this were just a story about me, one might write it off as a unique incident. But I see it all the time. Consider a story of a child whom I'll call Adam. When Adam was in elementary school, he changed from one school to another. Because the new teacher needed to place him in a reading group, school administrators gave him a reading test his first day in his new school, in much the same way that many colleges give language or math or other placement tests just as students arrive. At the risk of stressing the obvious, the first few days at a new school or on a new campus are not ideal days for giving tests. The students tend to be overwhelmed by the challenge of adapting to the unfamiliar environment, and their minds often are not on the tests they are given.

In Adam's case, he bombed the reading test, which meant that his school had a decision to make. It could follow the lead of his former school, which was essentially identical in quality to his new school, and put him in the top reading group. Or it could follow the superstition that a test tells all, and put him in the bottom group. The school put him in the bottom group, ignoring his former successful placement. He was being set up to fail.

After a few weeks, his teacher noticed that Adam was reading at a higher level than his classmates, and recommended that he be placed in a more advanced reading group. But the school, locked into the superstition that tests tell all, decided to ignore the recommendation and instead retest him. His reading test score was now at the level of the middle group, so the school put him in that group. Soon the teacher noticed that Adam was performing better than students in the middle reading group, so the
school once again gave him the reading test. This time he scored at the level of the top group. Given the school's ardent belief in the superstition that the test tells all, one would expect the school then to place Adam in the top reading group. The school didn't. It left him in the middle reading group.

Adam's mother and father made an appointment with very high-level personnel in the school to discuss why he was being retained in the middle reading group. The teacher was there, as was the principal, the reading specialist, and a school psychologist. The school officials explained to the parents that although Adam had scored at the level of the top group, he was now a full book behind the students in that top group. If they were to move him to the top group, he would lack the skills that the other students had. The parents were incredulous. They reminded the school officials that he was behind because the school had initially placed him in the bottom group.

The parents offered to help Adam with his reading. If he brought the book home, they would work with him to help him catch up in his reading skills. The parents felt qualified to assist him; both were PhDs and worked in education. But the school explained that its policy was not to allow reading books to go home, and that therefore they could not help. The school staff members were determined to ensure that the self-fulfilling prophecy they created would come true.

One would like to believe that colleges and universities are above such superstitions, but they are not. Rather, superstitions are firmly entrenched in our entire educational system, from elementary school onward. Colleges and universities sometimes use SATs and related tests in much the same way that elementary and secondary schools use standardized achievement tests. Sometimes the logic used is laughable—or would be, if real students' futures weren't at stake.

At one point, when I was working for The Psychological Corporation, a testing company, a case was brought to our attention as a result of a complaint. A student was applying to a graduate
program that, at the time, required a score of 25 on the Miller Analogies Test. This in itself was strange, because at the time the test had 100 items, each with four multiple-choice options, so 25 was the score one could expect to receive if one answered the questions at random. A student applied with a score that was below 25; she was unlucky in her responses. But the school admitted her anyway because she had other credentials that were excellent. She went through the program and was about to graduate with honors when she was handed an unwelcome surprise: she was told that, in order to graduate, she would have to retake the Miller Analogies Test and receive a score of at least 25 so she could meet the admissions requirement. Here the predictor—the test—was being given more weight than the criterion it was supposed to predict—university-level academic performance. Her complaint was that she should not have to retake the test. (The complaint later became moot when she took the test and received a score of 26. She graduated with her honors.) Such a practice may be atypical, but I have heard similar stories.

Academic superstitions do not have to be about test scores. In the case of a little girl whom I’ll call Margaret, it can have to do with other measures that an authority figure decides are important. When Margaret finished kindergarten, her Montessori School teacher wanted to hold her back because the teacher was convinced that Margaret had not yet developed the social skills she would need to succeed in first grade. Margaret’s parents persuaded the teacher otherwise and the teacher kindly promoted Margaret, who went on years later to study at an Ivy League college.

Some readers might wonder what all the fuss is about. After all, test scores do, on average, predict academic achievement, so why shouldn’t schools use them for this purpose? I have no objection at all to tests used in this way. Indeed, testing can be useful in diagnosing students’ strengths and weaknesses and in helping the students themselves to capitalize on their strengths and to improve in their areas of weakness.
The problem is that tests often are not used in this way. Some schools have cutoffs. These cutoffs may be explicit, in which case they are publicized; or they may be implicit, whereby college officials know they exist but do not talk about them. For example, in some business schools, if one’s GMAT score is below 600, one’s chances of getting in are reduced to practically zero. Many undergraduate as well as graduate schools use formulas for determining admissions. The assumption is that such formulas will accurately predict, at some level, who will succeed and who will fail. And it is further assumed that using such formulas somehow makes the process “fair.” Given the correlation of test scores with family income, it would be about as fair to plug family income into the equation instead of test scores.

**Financial Considerations**

This may sound ridiculous, but many schools—in fact, the large majority of schools—implicitly do include family income in the admissions equation. Those schools that are not “need-blind”—that is, which admit students in part based on financial-aid considerations—do take into account family income in making their decisions, whether formulaically or otherwise. For the most part, these schools would like to be need-blind: they just do not have the economic resources to reach their goal any other way.

Jesse Rothstein, an economist at Princeton University, has suggested that test scores essentially “launder” students’ socioeconomic background. Much of a student’s SAT score is attributable to the quality of the high school that a student attends, which in turn is a function of socioeconomic class. So test scores essentially serve as an effective substitute, however subtly disguised, for one’s family background on a student’s record. For even though admissions officers try to take extenuating factors such as lower socioeconomic background into account when considering scores, neither they nor anyone else knows the full extent of this relationship.

The concept of socioeconomic status is not perfectly defined.
It refers to one’s social standing in a way that takes into account social variables such as where one lives, economic variables like the wealth and income of the family, and educational variables such as educational background.

There are different views on why socioeconomic status correlates rather highly with test scores. The theories are not necessarily mutually exclusive. One theory is that children who grow up in higher socioeconomic circumstances are given more opportunities to learn the skills measured by the tests, and thus are in a better position to test well. A second theory is that the correlation is actually genetically determined. The idea is that smarter people pass on genes to their children that make them smarter; these genes lead them to excel and to reach higher socioeconomic levels; and they then pass on these “success” genes to the next generation, continuing the cycle. Most likely, genetic and environmental factors interact, as is the case with many aspects of human behavior.

CLOSED SYSTEMS, CLOSED DOORS

Standardized “aptitude” tests can, to some extent, create a closed society. They predict achievement because, to a large extent, they are achievement tests. The SAT Reasoning Test and the SAT subject tests (also called “SAT II’s”) are, in many cases, so similar that one cannot always distinguish them by their questions. Indeed, Binet’s intelligence tests, created at the beginning of the twentieth century, were designed to predict school achievement, so they were created essentially as achievement tests for skills that students should have acquired by a particular age or grade. Thus tests create a closed system: ability test scores predict achievement test scores because they are achievement tests. And the allegedly high “validity” of the tests makes users happy, because they essentially predict scores on other tests like themselves but are labeled something else.

At different points in time, societies, including our own, have found different bases for social stratification. In earlier times in
the United States, socioeconomic class was the primary criterion used to stratify students, for college admissions and earlier. Then test scores came to be used, and they seemed to be more “merit-based,” except that they produced largely the same results as the socioeconomic class indicators—the results were just different enough to give the appearance that wholly different measures were being used. At various times, admissions officers have used gender, socially defined race, and caste as bases for admissions. They have then found, to their satisfaction, that the criteria they used predicted future outcomes. This was not a kind of thinking particular to admissions officers. They were reflecting the societal context in which they lived. Remember that not so long ago, very few whites would have even considered the possibility of an African American becoming president of the United States. In most countries around the world, electing such a leader still seems impossible.

Imagine that we, as a society, were still using medical tests that were largely the same as those used at the beginning of the twentieth century. Would you be concerned? If you have ever had anything beyond a common cold, you would have reason to be. We should be equally alarmed that, as a society, we are using tests of abilities and achievements that are roughly the same as those devised a hundred years ago.\(^{15}\) Although our current ability and achievement tests are more refined, more sophisticated, and more carefully constructed, the deep structure of these tools remains unchanged. A narrow conception of intelligence seems to be prevalent in today’s society because of what is referred to as a closed system. A closed system is self-contained, internally consistent, and difficult to escape. A closed system, once it is in place, becomes self-perpetuating and difficult to change.

The vicious circle created by such a system gave rise to *The Bell Curve* by Richard Herrnstein and Charles Murray, a book that looks at the history of intelligence and class structure in the United States.\(^{16}\) According to Herrnstein and Murray’s 1994 analysis, conventional tests of intelligence, on average, account
for about 10 percent of the variation in various kinds of real-world outcomes. Although this percentage is not trivial, it is not particularly large either, and one then might wonder what all the fuss is about regarding the use of the tests. Of course, one might argue that Herrnstein and Murray have underestimated the percentage, but given their enthusiastic support for conventional tests, it seems unlikely they would underestimate their value. In fact, they may have overestimated the value of the tests for predictive purposes.

Clearly the tests have some value. But how much? In their book, Herrnstein and Murray refer to an "invisible hand of nature" that guides events so that people with high IQs tend to rise toward the top socioeconomic stratum of a society and people with low IQs tend to fall toward the bottom stratum. They present data to support their argument, and indeed it seems likely that, although many aspects of their data may be arguable, in U.S. society their argument holds true. For example, on average, lawyers and doctors probably have higher IQs than do street cleaners.

The problem is that although the data are probably correct, the theory behind the data is probably not. U.S. society is not as it is because of an invisible hand of nature, but rather because a closed system has been created. The United States and some other societies have created cultures in which test scores matter profoundly. High test scores are needed for placement in higher tracks in elementary and secondary schools. They are needed for admission to selective undergraduate programs. They are needed again for admission to selective graduate and professional programs. It is really quite difficult to imagine how a person could gain access to many of the highest-paying and most prestigious jobs if he or she did not test well.

This system is tragic in many ways. Students spend years in primary and secondary school working hard (or not so hard) to achieve their goals, and then do the same in college. Sometime during their career, they take a test that lasts three to four hours.
Then that test score ends up having a weight equal to those products of years of effort and dedication in the admissions decisions of college and graduate programs. If test scores were as valid as other achievements in predicting success, perhaps this system would make sense. But they are not.

Instead, the 10 percent figure of Herrnstein and Murray implies that IQ-like abilities matter some, but not much, for life success. Other abilities that are important for life may not be measured effectively by standardized tests. For example, the creative and practical skills that matter to success on the job typically are not measured on tests used to get into school. Society may be overvaluing a fairly narrow range of skills, even if that range of skills may not serve individuals particularly well on the job.

In a reductio ad absurdum of the whole IQ-testing mentality, the New London, Connecticut, police force rejected the applicant Robert Jordan for having too high an IQ.\(^{17}\) His score, 125, though far from stellar, was well above average. The police force argued that someone with too high an IQ might get bored quickly and then decide to leave the police force, thereby wasting the city’s money spent in training. Some might find it disconcerting that a police force would value mediocrity, but apparently, this one did. The potential policeman sued and lost.

It is scarcely surprising that ability tests predict school grades, because, again, the tests were originally designed explicitly for this purpose. This historical fact makes it obvious how the United States and some other societies have created closed systems. Certain abilities are valued in instruction, such as memory and analytical abilities. Ability tests are then created that measure these abilities and, thus, predict school performance. Then assessments of achievement are designed that also assess these abilities. Thus it’s little wonder that ability tests are more predictive of school achievement than of success in the workplace. Within the closed system of the school, a narrow range of abilities leads to success on ability tests, in instruction, and
on achievement tests. But these same abilities are less important later in life, for instance, on the job, when providing for a family, or when contributing to a wider community. The costs of this closed system are high. Closed systems seal off individual options and distort society, depriving many individuals of opportunities they should have. Society is also deprived of their talents. So why did such a closed system emerge?

In addition to the reasons for the development of closed systems mentioned earlier, there may be a unique historical explanation for why the tests we use are so narrow. An early intelligence theorist, Charles Spearman, active in the first decades of the twentieth century, believed that all there is to intelligence is what he called “general intelligence,” or g. Belief in this general factor remains entrenched among many psychologists, even today.¹⁸

A BROADER UNDERSTANDING OF INTELLIGENCE

Many others, however, believe that there is more to intelligence than just one general ability.¹⁹ For example, Howard Gardner has proposed eight “multiple intelligences”—linguistic (understanding what you read), logical-mathematical (balancing your checkbook), spatial (using a map to find your way around a new city), musical (keeping tune in singing a song), naturalist (observing and understanding patterns in the natural world), bodily-kinesthetic (kicking a goal in soccer), interpersonal (understanding other people), and intrapersonal (understanding yourself).²⁰

A theory such as Gardner’s suggests that the tests we use for college admissions are too narrow, because they primarily assess only linguistic and logical-mathematical skills.

Furthermore, depending on the situation, attributes that have nothing to do with intelligence (at least, according to the present and most other definitions) can end up becoming conflated with intelligence, thereby skewing any test results. For example, I attended classes in a number of typical one-room elementary schools in Jamaica. There was no barrier separating the many
classes in the single room, so the noise level was constantly high. I found myself wondering what Binet might have put on his intelligence test if he had wanted to calibrate it to measure future success for these students, and concluded that he might have decided to include a battery of hearing tests. One’s ability to hear the instruction and the test questions, both of which were typically given verbally, was critical. In this situation, those who did not hear well, perhaps simply because they did not sit in the front center of the classroom, fared worse than those who could hear well.

The importance of hearing to measures of intelligence is not just hypothetical. When I mentioned my observation in a colloquium, an individual from Guyana commented that she had grown up in similar schools and had always wondered why the smartest students sat in the front of the class. In this case, sitting in front, where it was easiest to hear and to be engaged, may well have made students appear smart. The teacher probably did not think that good auditory (sensory hearing) abilities were a component of intelligence, but he or she might have easily conflated the effects of such abilities with intelligence. Similarly, students with poor vision who do not have the benefit of corrective lenses may appear to teachers and school staff to be not very bright.

The experience in Jamaica also points out one other important fact, namely, that much research on intelligence is based on the assumption that all students have an equal chance to succeed on ability tests and in school. In fact, they do not. For example, in a study done in Jamaica in 1997, my colleagues and I studied the effects of intestinal parasitic infections (most often, whipworm) on students’ cognitive functioning. Our study showed that infected students tended to do worse than other students on tests of higher-order cognitive abilities, even after controlling for possible confounding variables such as socioeconomic class. It also revealed that, although antiparasitic medication improved physical health, it had no effect on cognitive-ability test scores. Presumably, the deficits caused by the infection had built up over
many years and were not alleviated by a quick-fix pill. Students who are parasitically infected find it hard to concentrate on their schoolwork because they do not feel well. The data showed that the cumulative effect of missing much of what happens in school probably cannot be reversed quickly. Indeed, students in all societies who suffer from health problems, including poor nutrition, or who feel unsafe at home or school, do not have equal chances to succeed.

Societies can and do use a variety of criteria to sort people. Some societies use caste systems—whether explicitly, as in India, or implicitly, as in the United States. Others use or have used race, religion, or parental wealth as a basis for sorting. Many societies combine criteria. After a system is in place, those who gain access to the power structure, whether through elite education or another means, are likely to look for others similar to themselves to place in positions of power—simply because there probably is no more compelling basis of interpersonal attraction than similarity. The result is a potentially endlessly looping closed system.

Why have we reached this point? First, the cart somehow came to be placed before the horse: commercial interests came into play before the science of assessment had much opportunity to develop. Second, a small number of commercial enterprises largely monopolize the testing business, and many of the researchers in the field (including myself) have either worked for, or have been supported financially or otherwise by, these organizations. Such organizations, like any others, will continue doing what is profitable for them. Third, we have been and continue to be entranced by the notion of accountability, while largely ignoring just how narrow our cognitive and educational measures of accountability are. Fourth, the tests have seemed roughly consistent with the notion of a general ability that pervades many psychological performances. And finally, by some measures, the tests have worked: they predict various kinds of success at some modest to moderate level. The question, then, is not whether
these tests work at all, but rather, how we can improve the overall process of student assessment.

It is tempting to blame the "evil" testing companies for the problems we have in admissions testing. But they are no better or worse than the large majority of other companies. They are there to make money. (Even the nonprofits are there to make money, although they have a different accounting system.) If customers—the colleges and universities—were to insist on other types of tests, the companies would create them. But the colleges and universities are comfortable, for the most part, with what they have. Consequently, there is no one entity that is to blame for what we have today. Instead, we have a case of an inadequate system that has become entrenched because it benefits many powerful people and because, to date, the voices of those who do not benefit have not been forceful or persuasive enough to enact change.
HOW WE GOT HERE:
THE TRADITIONAL
COLLEGE APPLICATION

What are the elements of a traditional college application? Which assessments are used, and how are they combined and weighted to arrive at an admissions decision? An essential tenet of educational and psychological testing is to use assessments in combination in order both to capitalize on their strengths, and, by averaging out possible errors, to compensate for their weaknesses.

HIGH-SCHOOL GRADE-POINT AVERAGE

Grade-point average (GPA) still is the most important factor in admission to most colleges and universities. It has many strengths as an assessment.

First, the best predictor of future behavior is generally past behavior of the same kind. For example, if someone has tended to overeat during holidays in the past, they probably will in the future as well. If someone has been charitable in the past, he or she probably will be in the future, too. And if someone has earned good grades in the past, he or she probably will do so again in the future. Because academic work forms a cornerstone of college education, and because failing academic work can cause a student to drop out early, it makes sense that college admissions offices would rely on high-school GPA as a fundamental basis for making their decisions.
A second strength is that high-school GPA, although a single number, actually represents the level of mastery of a wide range of skills. It reflects a student’s academic ability, of course, because getting good grades requires students to acquire a knowledge base and then to reason with it. It further reflects the student’s practical abilities: can the student understand what is expected of him or her, devise strategies for studying for different types of tests (such as multiple choice and essay), figure out what the teacher is likely to test, and budget the available time in a way that allows him or her to excel in not just one subject, but several? Moreover, high-school GPA reflects, as much as or more than anything else, motivation—the student’s willingness to work hard to achieve high marks in school.

Third, GPAs are readily available. Nearly all schools calculate a GPA for their students, so the admissions office need not require anything extra of the student in order to get the information, except perhaps the student’s permission to have it released.

The GPA, however, also poses challenges. First, high schools differ in their quality and in the quality of the students that attend them. In some high schools, having a 3.8 average may mean that the student has completed a highly rigorous selection of courses and has done extremely well in them. In other high schools, courses may be much weaker and the grades in them mean much less. Even within high schools, courses differ widely in how academically rigorous they are.

Second, high schools differ in what their grades mean. Long ago, “C” meant “average.” In theory, roughly 8 percent of grades were A’s, 24 percent were B’s, 36 percent were C’s, 24 percent were D’s, and 8 percent were F’s. Schools may have differed in the exact percentages used, but the term “gentleman’s C” referred to the grade received by a student who did not work particularly hard and who made no effort to achieve at a level above the average. Today, by contrast, a C is often considered a disgrace. In fact, I have had any number of students complain about
grades of A−. They want to be the best and anything less just will not do.

Third, the desire for high grades is partly a result of high motivation, but it also has been fueled by grade inflation. An A just does not mean as much anymore. Neither does a 4.0 average. With corrections introduced by schools for course difficulty and other factors, a 4.0 average may place a student in the upper portion of a high-school class, but nowhere near the top. Efforts are made every now and then to hold grade inflation in check, but such efforts have not been particularly successful.

Fourth, high schools differ in what courses even count toward the GPA. In one school, a course such as woodworking may not count at all; in another school it may count but be weighted only lightly; in another school, it may count the same as any other course. The school may indicate its methodology, but admissions officers typically are not going to recompute the whole GPA to give more value to certain courses.

Fifth, grades represent in some degree the skill of knowing what teachers want and giving it to them. Of course, students will have to do the same in college, but will they use the same skill? In a rigorous high school, the expectations of the teachers may correspond to those of college instructors—for example, students may be asked to think independently, debate appropriately with others, and conduct in-depth outside research. But in a less rigorous school, an A may mean simply that the student was well behaved and did exactly what he or she was told.

Some schools are familiar to admissions officers. When I worked in admissions at Yale some time ago, I felt I could interpret quite well the grades and GPAs from schools with which I was familiar—usually those that consistently sent a number of students to Yale over the years. But I also realized that my knowledge could bias me toward candidates from those schools.

After I was admitted to Yale, I went to work in the college's Office of Institutional Research under a psychologist, now de-
ceased, who computed predicted grade-point averages. He was someone who even then, in 1968, represented a past generation of psychologists. He predicted GPAs using a hand calculator (and when computers came into use, and he found himself without a job, to my astonishment he took to checking the computer's calculations by hand). I discovered that he used a correction factor that took into account the high school one attended. Yale had been using percentage grades and the psychologist in charge would convert a student's grades in high school to expected percentage grades at Yale. In my case, for example, because I had attended a public high school that was largely unknown to him, he had automatically subtracted nine points. That meant that the highest predicted GPA I could attain was a ninety-one.

The psychologist in charge was only trying to do his job, albeit in an outdated way. But the problem that he was addressing is still not totally solved. When an admissions officer sees a GPA from an unknown school in rural Appalachia, or an inner-city high school known more for its disciplinary problems than for the academic achievement of its students, the admissions officer may make the same calculation in his or her head, perhaps not even quite consciously, that the psychologist had made explicitly. Schools that seek diversity may end up admitting some students from the unknown high schools. But only a few.

The problem of underrepresented schools has another origin: for these students, college may not seem like a viable option, and a selective college may not be something they even consider. This gap is, in part, due to inadequate resources for college guidance at the high schools. In addition, guidance counselors in some remote schools may not even be aware that their students have a chance of attending colleges that, to the counselors, seem out of reach.

The irony is that going to a more selective school does not necessarily mean that one will have to struggle more to achieve good grades. At the university level, for example, Harvard,
among the most selective colleges in the United States, has been as wracked by grade inflation as anyplace else, and perhaps much more so than many of the community colleges with which it would never view itself as competing. Grade inflation has affected many high schools and private preparatory schools as well, including some of the most selective ones. The general point is that grades are hard to interpret, for a wide variety of reasons. Fortunately, admissions officers look for more.

**CLASS RANK**

Most high schools provide a measure of a student’s rank in class—that is, how far he or she is from the top of the class. The information in class rank is partially but not wholly duplicative of the information in GPA. This is because class rank controls for the severity of the grading in the school. In one school, a B (3.0) average may place a student in the upper third of the class, whereas in another school, it may place a student well into the bottom half. The class rank partially adjusts for these differences in grading.

The class rank is always interpreted in the context of how many students are in the class, and is usually interpreted in the context of the school one attends. A class rank of one means a different thing at a highly competitive school than at a small rural school with students who are much less well-equipped academically. But school quality is not always factored in: for example, being in the top 10 percent of your high-school class in Texas guarantees admission to the University of Texas, regardless of the school you attend.¹

Some high schools and private preparatory schools prefer not to rank their students. In fact, if there is a trend in recent times, it is toward high schools’ not computing class rank. The advantage of such a decision is that it takes some pressure off students during their high-school careers. The disadvantage is that college admissions officers find it harder to evaluate applicants from such schools, so students could potentially be put at a disadvantage. If
the school is known to the admissions officers, however, they may calculate, even implicitly, a rough class rank based on what they know about the school.

**COURSE LOAD AND PROFILE**

College admissions officers consider not just the numerical value of the GPA, but also the courses that make up that GPA. The course load and profile can tell the admissions officer quite a bit about the student’s academic skills and motivation. Is the student taking very challenging courses or relatively easy ones? If a foreign language is expected by the college, how many foreign-language courses has the student taken, and how many languages has she or he studied? Has the student taken the requisite mathematics courses needed for college success? Are there special courses on the transcript, such as honors or Advanced Placement (AP) courses, if they are available in the student’s secondary school? Is the student in a special program known for its rigor, such as the International Baccalaureate (IB) program? What kinds of electives has the student chosen? Information about the course load and profile can reveal information not available in the composite GPA.

At the same time, the admissions officer needs to factor in extenuating circumstances. Not all schools offer AP courses or the IB program; and even if they do offer AP courses, their offerings may be limited. Schools may also have various restrictions on who can take such courses. For example, in one school, AP courses may be available to virtually anyone, whereas in another school the courses may be available only to the very best students. So the admissions officer needs to factor in not just which courses were taken, but which were available to be taken. The admissions officer also needs to see the course profile in the context of the students’ other activities. For example, students who are very involved in high-school sports may choose courses that will allow them to succeed academically even though they have to spend a lot of time on the playing field. The officer may or may
not choose to view such extracurricular activities as compensatory for taking a lighter course load.

STANDARDIZED TEST SCORES

Much of this book is about standardized tests, so I will not repeat here what is covered elsewhere. In summary, there are two standardized tests that have a monopoly in the college-admissions market in the United States—the SAT and the ACT.

Consider some of the history. The College Board was created in 1901, and the first of its tests was given in that year. At that point, the tests involved essays measuring achievement in school subjects such as history, Latin, and physics. But the enterprise as we know it today did not get under way until June 23, 1926, when the first Scholastic Aptitude Test (SAT) was administered. The test was created by a Princeton psychologist, Carl Brigham, and it had sections related to antonyms, arithmetic, analogies, and paragraph reading, among other things. Administered to roughly eight thousand students, the test lasted only an hour and a half, despite having more than three hundred questions. Significantly, the kinds of questions used then did not differ much from those used today, and the underlying skills that the tests measure have scarcely changed.

The name “Scholastic Aptitude Test” was later dropped in favor of “Scholastic Assessment Test,” and this name in turn was later dropped in favor of simply using the acronym “SAT” without its meaning anything in particular. Different reasons might be given for this shift, but the most obvious one is that neither the College Board nor anyone else is quite sure what the test measures, because it is not based on any particular scientific theory of a psychological construct but rather on a pragmatic assessment of what will predict scholastic success in college.

The SAT Reasoning Test has three parts, now called critical reading (formerly the “verbal” section), mathematics, and writing. The critical reading section contains two types of items, sentence completion and passage-based reading. Sentence comple-
tion measures the student’s word knowledge and understanding of how words fit together to form a comprehensible sentence. An example might be “The Parliament ______ the bill and, as a result, it ______ the law, resulting in a steep increase in the taxes citizens paid to the Crown.” Options might be (a) rejected . . . enacted, (b) passed . . . became, (c) discussed . . . accepted, (d) disputed . . . nullified. Passage-based reading, by contrast, measures one’s ability to read a long passage and understand and analyze it. And the critical reading section assesses the student’s understanding of vocabulary in context, literal comprehension, and ability to engage in extended reasoning that goes beyond simply remembering what one has read. In some sections, test takers may be asked to compare two passages. Content generally covers the natural sciences, social sciences, humanities, the arts, and personal narratives. The SAT critical reading test is scored on a scale from 200 to 800.

The SAT writing test has two sections, one of which measures grammar, English usage, and choice of words in contexts, and the other of which is an essay that requires the student to take a point of view and defend it. The multiple-choice section is scored on a scale from 200 to 800, and might show a sentence or paragraph and ask how it could be improved, or ask whether there is a grammatical error in a sentence. The essay is scored in a largely formulaic way. It is not designed to measure creativity in writing but rather the mechanical skills of knowing how to set up an argument, carry it through, and draw it to a conclusion. Thus, for better or worse, it is highly trainable. In a sample essay on the College Board website, students are asked whether memories help or hinder one’s ability to learn from the past and succeed in the present. Students have just twenty-five minutes to write their essays.

According to the College Board website, an ideal essay develops a point of view on the issue at hand, in this case, the role of memories in learning from the past. It shows excellent critical thinking, and uses evidence and examples that help to bolster
the argument being made. The essay must be exceptionally well organized, focused, and systematic in its development of an argument, and needs to include a variety of sentence structures. Finally, it should show nearly flawless use of words, grammar, and syntax. It is scored on a scale from one to six by each of two readers.

The SAT mathematics section has three parts that assess the student's knowledge of numerical operations, algebra, geometry, probability and statistics, and data analysis (for example, understanding and interpreting a graph). The highest level of difficulty tests material that would be covered in the third year of college-preparatory mathematics. The test has both multiple-choice and free-response items. Students are encouraged to use a calculator, and the test is not intended to measure recall of basic formulas; indeed, these basic formulas are provided. The test is scored on a scale from 200 to 800.

The second of the major college admissions tests, which tends to be used more in the Midwest and the South of the United States, is the ACT, which traditionally stood for American College Test but, like the SAT, now goes just by its acronym. The ACT was created as a competitor to the SAT by Everett Lindquist, and was first used in the fall of 1959. It supposedly measures achievement more directly than does the SAT, but in fact the SAT subject matter tests are even more direct measures of achievement than is the ACT. The ACT, traditionally, has consisted of four subtests: English, mathematics, reading, and science reasoning. It also now includes an optional writing test. Scores range from 1 to 36 on each of the subtests as well as for the overall composite score, which is an average of the other tests. If the test taker takes the writing test, it is not included in the composite.

The English test requires forty-five minutes and covers English usage, similar to the multiple-choice part of the SAT writing test. The mathematics test lasts sixty minutes and includes algebra, plane geometry, coordinate geometry, and trigonometry questions. Calculators can be used. The reading test lasts thirty-five
minutes and measures reading comprehension for four passages, including fiction, social science, humanities, and natural sciences. The science reasoning test lasts thirty-five minutes and requires the test taker to understand scientific writing and representation of scientific data. The writing section is thirty minutes long and involves a social issue that is relevant to high-school students.

It is easy to see why the SAT and the ACT are popular among college admissions officers and other administrators and faculty. The tests have several positive features. First, the tests measure skills that are relevant to college success. Students will need to read a tremendous amount of material in college, and they will need to understand and analyze it. Moreover, the material they will read will be in a variety of fields like those sampled on both the SAT and the ACT. College students will also have to know high-school mathematics in order to succeed in college math and science courses. They will need to be able to write well in order to complete essays on examinations as well as term papers and the like.

Second, the tests seem to provide a common metric across different students and schools. An SAT score of, say, 600 refers to the same number of correctly answered questions (after correcting for guessing) in one school as in another. In contrast, the GPA may mean very different things in different places.

I say “seem” to provide a common metric because a given score does not really mean exactly the same thing for each student who achieves it. One student may have grown up as a native English speaker with a lot of books in the home and highly educated parents who were able to educate the child in a way that would help maximize scores on the tests. Another student may have grown up with a foreign language as a native language, or even with parents who do not speak English at all. Another student may have had few books in the home and parents who were not able to provide an enriched educational experience. Admissions officers are trained to take such differences into account,
but no one can do so perfectly because no one knows exactly what the effects of different upbringings are. Moreover, the first student likely will be at an advantage in college by having mastered the skills to achieve high grades. So the admissions officer must decide just to what extent he or she wants to take the background differences into account.

Third, the college does not have to pay for the test—the student applicants do. And finally, the tests give the appearance of being objective, so that colleges do not have to worry that the results are a reflection of who happens to be scoring them. The possible exceptions to this are the results of the writing tests, but even those are scored in ways that deliberately overlook novel, imaginative responses in the search for elements that can be measured somewhat objectively.

With all these advantages, why would anyone in his or her right mind not weigh the SAT or the ACT heavily (although certainly not exclusively) in admissions decisions? Well, for one thing, although the tests measure skills relevant to college success, they do not measure all such skills. Students need practical, commonsense skills to know how to study and organize their time. They need metacognitive or self-understanding skills to recognize their own strengths and weaknesses and how they might improve. They need motivational skills and attitudes to get themselves to work hard. And they need creative skills to generate novel ideas.

In addition, exactly how common is the metric provided by standardized test scores? I mentioned earlier that scores can mean different things for different individuals. Can the meaning of a given score vary by group as well?

In our studies of various ethnic groups living in San Jose, California, Lynn Okagaki and I found that each group had a rather different idea of what it means to be intelligent. Teachers tended to reward those children who were socialized into a view of intelligence that happened to correspond to the teachers’ own.8

In any culture, practical intellectual skills matter for predict-
ing adaptation to everyday environments. But a child growing up in inner-city Detroit, generally a low-income part of Michigan, faces challenges different from those of a child growing up in Grosse Pointe Farms, a generally wealthy suburb of Detroit, and parents socialize their children to develop the skills they need to face the challenges of their own environments. So the SAT actually may measure somewhat different skills for different children, depending on which skills the parents emphasized as the children were growing up. For example, the SAT mathematical test measures different skills for someone who has learned algebra and geometry than for someone who has to figure out the answers without that knowledge.

It is also important to realize that although the college does not pay for standardized testing such as the SAT or the ACT, students do. If we add up the costs of multiple test-takings, of books that parents buy for their children to help them prepare for the tests, and of courses or tutors, the costs become non-trivial. Moreover, there is also the “opportunity cost” to students who attend high schools that integrate what amounts to SAT preparation into their curricula. That is, the time the students spend preparing for these tests is time they are not spending doing other things that might be more worthwhile, such as engaging in athletics, creating art projects, studying music, or even reading literature instead of memorizing vocabulary words.

The ACT and, more recently, the SAT, offer students score choice—that is, they can choose which scores to send to colleges. But score choice poses a bit of a dilemma. Is a student who receives a 550 and a 650 on, say, the math SAT, really better than a student who twice receives a 600? The student who received the disparate scores can report only the 650 and thus look somewhat more skilled mathematically than the student who received the 600 twice. But is the student really more skilled? The average scores of both students are the same. One interpretation, of course, is that the 550 merely represented a “bad day.” But error of measurement goes both ways—it can result in scores that are
too low or too high. So score choice may skew the information available to colleges. It also poses another dilemma. If the student need only report highest scores, it encourages students to keep retaking the test—and paying for it—in the hope that he or she will hit the jackpot, securing a score that is appreciably higher than the other ones he or she has received.

Do we overvalue objectivity in tests? Consider the oft-told story of a man who is frantically looking at night for his lost keys. A police officer sees him looking and offers to help. Fortunately, the area where the man is looking is the one place that is well illuminated. The police officer and the man look for quite a while, with no success. Finally the police officer asks, “Are you sure you lost the key here?” To which the man responds: “Why, no, I lost them over there. But it is pitch black there, so I’m looking here where the light is better.”

In standardized testing, we have tended to measure what is easy to measure. Arguably, this seemed right and proper in the 1920s, because testing was just starting out and it was impressive, at that point, that the testers could measure anything in a standardized and useful way. But the years passed and little changed. In testing, we have looked where the light is best, but in doing so, we have ignored what may well be the most important keys.

The so-called objective tests do serve a purpose in college admissions. But there is no need to limit ourselves to such tests. The best assessments use a variety of different assessment methods. Ideally, college admissions officers would combine the traditional objective assessments with others that, although more subjective in their scoring, allow students better to show the range of skills they possess.

Testing companies like the Educational Testing Service (ETS) and ACT are moving ahead with several projects to measure some of these other skills, such as the ETS Personal Potential Index, which seeks to measure skills beyond what the Graduate Record Examination (GRE) measures. One would hope that
more and more assessments would move in this direction, and that they would become a more important part of admissions processes, for undergraduate as well as graduate admissions.

THE COMMON APPLICATION

Before the mid-1990s, each college had its own separate application, and students had to do a lot of repetitive work, filling out the same information again and again on multiple applications. Today, the Common Application is used by many colleges to save applicants this trouble.\(^9\)

The Common Application asks for personal and family demographic information; test scores; lists of athletic or other honors and academic distinctions; principal extracurricular, volunteer, and family activities; paid work experience, if any; and future plans.\(^10\) It also gives the applicant an opportunity, in 150 words, to elaborate on one of his or her activities. A longer, 250-word essay can be either open-ended, with the applicant choosing what he or she wishes to write about, or based on a prompt, such as to write about an experience one has had or a viewpoint to which one subscribes. Because the Common Application serves such diverse and heterogeneous applicants and colleges, the essay questions are pretty much “one size fits all.” The application also asks about disciplinary history and any additional information the applicant may wish to provide.

Forms are included with the application that are designed to be given to one’s teachers and to the guidance counselor (or equivalent). The teacher form asks about attributes such as academic achievement, intellectual promise, maturity, motivation, leadership, and integrity. It also asks for an overall rating. The guidance-counselor form includes a question about how class rank is computed, as well as a request for academic, personal, and overall evaluations of the applicant.

The Common Application does a good job of organizing a wide variety of information about the applicant in a way that can be used by many different colleges and universities. Moreover, it
cuts down enormously on the work that applicants need to do in order to apply to multiple schools.

Yet there are some problems with the Common Application, as there probably would be for any application of its kind. First, as mentioned earlier, the essays have long been fairly standard and open-ended. Consequently, there has been an opportunity for unscrupulous merchants to build up banks of successful essays and sell them to students applying to colleges. There is little one cannot buy on the Internet these days. At best, students may use such essays as models. At worst, they may recycle old essays and use them as their own. The extent to which this dishonest practice occurs is unknown.

A second feature of the Common Application is that the essay topics do not particularly encourage creative responses (nor are they even designed to do so). Rather, they encourage students to spend years “building up records” that then can translate into college admissions success. As a result, the experience in high school and even earlier can become largely about building a record for college. In some cities, such as New York, the competition starts early, with children vying for places in kindergartens whose graduates tend to win placement in the city’s more prestigious private schools. At least in some circles, the college admissions game begins before students even enter elementary school.

A third feature of the Common Application is that recommendations, like grades, have become inflated over the years. Top-notch ratings of students by teachers and guidance counselors have become more common, making it harder for the colleges to distinguish among applicants. This difficulty is not limited to the Common Application, but neither does this application provide a remedy to this vexing problem.

A fourth feature is that although teachers are asked to give a number of distinct ratings for the student, the ratings in fact suffer from what is often known as a “halo effect.” This simply means that, if the teacher has a generally positive impres-
sion, most of the ratings turn out to be quite positive, and if the teacher has a generally negative impression, most of the ratings turn out to be less favorable, at least for a very selective college. Thus although it would be nice to receive distinct ratings for each of the rated elements, one rarely does. Rather, they reflect a general glow that the applicant emits in varying degrees. In the end, the written comments are often more helpful than the summary ratings.

A fifth feature is that the ratings can mean very different things in very different schools. The top student in Schnoginville may indeed be the best that the town has seen in the past ten years, but he or she may be competing with students who are not very competitive. The result, as discussed elsewhere, is that admissions officers may take more seriously recommendations from those schools and guidance counselors whom they know and trust.

INTERVIEWS
Interviewing practices differ widely from school to school. Some colleges do not offer interviews. Others do, but through alumni groups only. And still others offer on-campus interviews, but they may or may not use the interviews to evaluate candidates. Some years ago, when I worked in admissions at Yale, I did a cost-benefit analysis of the Yale Admissions Office interview.11 Somewhat to my surprise, I found that the interview really did not count very much in the admissions decisions. The admissions officers realized that the interview is of limited value. It can help spot a few outstanding candidates, and some real duds. But as I discovered myself when I conducted college interviews, most of the applicants look good but not terribly well distinguished from the rest. Curiously, however, when applicants were asked to evaluate how well they had done, they thought they did much better than they did. And perhaps because they thought they presented themselves so well, they had a very positive impression, overall,
of the interview experience. As a result, I recommended retaining the interview, primarily for public-relations purposes.

Alumni interviewers enable colleges to reach out to large numbers of applicants in ways that in-house interviews do not, but the range of talent among the interviewers is great, as is the range of standards to which they hold the applicants. It often is hard to know just what to make of the reports. Indeed, when I applied to Yale, my alumni interviewer was from another generation in which the values of the university had been very different. It was unlikely that he would have even known what the school was looking for in the modern age of 1967.

Interviews sometimes feel more valid than they are, because it is often difficult to separate self-presentational skills from the more basic attributes of the person. Self-presentational skills do matter in life, so having the interview may predict the kind of impression a person will make on others, at least early on. They will be a disadvantage, however, for young people who may be slow to warm up or who feel intimidated by the interview setting. Like everything else, the interview has its strengths and its drawbacks as a means of assessing an applicant’s characteristics.

EXTRACURRICULAR ACTIVITIES

Extracurricular activities form an important part of the application for many universities. They are considered important because they potentially reveal aspects of the applicant’s personality and character that would otherwise not come to light. Admissions officers often are interested in seeking future leaders and productive citizens, and they know that test scores and grades only tell one so much. Extracurricular activities might speak directly to other qualities.

Not all schools count extracurricular activities. Some schools, especially state ones, are much more interested in class rank or a combination of class rank and test scores. In most of Europe, extracurricular activities do not count at all. Whether to count
them depends as much on one's philosophy of admissions as anything else.

Even in the United States, colleges generally do not weigh extracurricular activities as much as they value test scores and grades. Why not, if such activities are among the few measures of character? Because self-reporting of extracurricular activities brings up several important questions, including:

- Are extracurricular activities relevant to whether a student will succeed academically? The first priority of the college or university is to ensure that the student can do the academic work at the institution. It is of no use to admit someone with a stellar extracurricular record who will flunk out after one semester. And in selective colleges, admissions officers seek students who will not merely pass their courses, but excel in them. Extracurricular activities generally do not say much about an applicant's ability to do the academic work of the school.

- Just how meaningful was the applicant's participation in the activity? It is hard to quantify extracurricular participation. Suppose, for example, the applicant reports having been president of the debating club. Questions then arise. Was he or she really president of the club? If so, how big a job was it? For example, how many members were there on the team, how much time did the student invest, and how many events were there? And what did the student actually do in the activity? Was he or she an active president who moved the club forward, or was the activity largely passive, serving to fill up a college application but little else? Ideally, the combination of sources of information will answer these questions, but not always. And although admissions officers can assign ratings to such activities, these ratings can seem subjective in comparison with test scores and grades.

- How many opportunities did the school provide? At some
schools, there may be room for everyone to be president of a club, whereas at others, there may be little opportunity for a student to shine in clubs or other after-school activities.

- Who wrote the description of activities? There is always the risk that the description of activities was written by someone other than the applicant, which would not only be inappropriate, but also would increase the chances that the description is inflated or flat-out wrong.

**LETTERS OF RECOMMENDATION**

Letters of recommendation provide an opportunity for teachers, guidance counselors, and occasionally others to provide information that may be obtained no other way. They can tell the college about a student's academic performance, abilities, initiative, curiosity, character, motivation, and the like—all from the viewpoint of someone who works directly with the student, often on a daily basis, and knows him or her well.

Sometimes, however, letters of recommendation are not nearly as useful as they could be. Why?

- Just as grades have been subject to inflation, so have letters of recommendation.
- Letter writers, even if assured the letters of recommendation are confidential, may worry that if they are less than positive they can be sued later. Their worry may or may not be justified. Although the confidentiality of letters is protected under the Buckley Amendment if the student signs an access waiver, there are no absolute guarantees. I was once burned myself when I wrote a supposedly confidential letter and then found out that the student about whom I had written had nevertheless learned the contents of my letter. Leaks do happen, even if only rarely.
- Some letter writers do not know the students as well as would be desired, especially if the writers are responsible
for very large numbers of students. Even if the letter writers know the students, they may not know about all of the attributes on which they are asked to comment.

- Although there are often multiple ratings that can be given to students, there can be a halo effect, which as explained earlier will cause all or almost all of the ratings to be similar. Halo effects are common not only in letters of recommendation, but also in many other kinds of ratings that people make.\(^{13}\)

- Some letter writers are very experienced in writing letters of recommendation to competitive colleges and know better how to get the result they intend. They are experts at coded messages that downplay or obscure negative attributes while accentuating the positive. Other letter writers, say from high schools that send only a few students on to selective colleges, may not understand the system, so their letters may not give the student this added benefit. Admissions officers, too, are inclined to weigh more letters from schools and counselors they know and respect, which may give an edge to students who come from schools that have sent multiple students to that college.

The bottom line is that letters, like all other indicators, have strengths and weaknesses and have to be considered in the context of all available evidence.

**TACIT FACTORS**

There are other factors that affect admission that are not specifically part of the admissions application. They are factors that may weigh in a decision, even heavily, but that are tacit rather than explicit. These factors may include things such as the applicant’s gender; ethnic affiliation; geographic origin; religion; specialized skills in areas like athletics, music, drama, or visual arts; and personal connections to the university, such as through par-
ents or other relatives who work at the university or are alumni or large donors.

Many colleges and universities will declare that they seek diversity among their students but will not elaborate beyond that statement, and with good reason. If they become too specific, they risk legal entanglement. Diversity is genuinely a good thing: you can learn things from someone whose upbringing was very different from yours that you could never learn from someone who grew up next door. But seeking diversity is rarely straightforward.

At a university that is, say, a major Division I competitor in varsity athletics, anyone with half a brain knows that athletes receive different consideration than do non-athletes. Schools differ in just how much extra consideration they will give applicants who are athletes, but most do weigh athletic and other skills in their assessments.

Different people will give different justifications for these special categories. Usually they are grouped together under "diversity," although sometimes preferences impede diversity as, for example, in the case of facilitated acceptance for children of people who are well connected. As I write, there has just been a nationally covered scandal at a major state university regarding a special list that existed for children of well-connected parents. What amazed me was not the scandal but the number of people who acted as though they did not realize what was going on. The truth is, what happened at this state university happens every spring, in one form or another, at many other colleges and universities. Such patronage risks perverting the whole purpose of college and university admissions—which is supposedly to admit the best candidates. But it is very common.

In the case of special talents, one might argue that particular applicants should be accepted because of their unique ability to contribute to the university. For example, a given athlete might not be a top academic performer, but he or she is good at some-
thing that matters to the college or university. That something happens to be perceived as increasing alumni interest and participation in the school and probably financial donations as well. University athletics are a big business. Wins enhance donations, sometimes greatly, as well as ticket sales, which in turn enhance university revenues. Is it right to admit students because they enhance university revenues? It depends on whom you ask. Personally, I think such policies, at least when taken to extremes, lead universities down paths they will come to regret. In Europe, athletic prowess counts for nothing in college admissions. But in the United States, where there are teams and the college or university has a stake in their success, it would be difficult to remove athletic considerations from the admissions procedures in some universities. Moreover, athletics provides one way to build leadership skills among students.

Pragmatically, the temptation to succumb to various pressures is enormous—even though it can be, and is, overcome. At one of the universities where I have worked, a problem emerged when the child of a major potential donor was being considered for admission but was unqualified. The admissions office was aware of the enormous stakes involved. And it did try to find reasons to admit the child. In the end, however, he was rejected. The university was willing to consider bending its standards a bit, but it wasn’t willing to break them. The donor was furious and said he would never donate again, and he hasn’t. So the decision to reject was probably costly to the school. There was only one justification for the negative decision, really—it was the right thing to do.

Most such cases, though, are not black and white but instead gray. Perhaps the first oboist has just graduated and the orchestra desperately needs another. Or perhaps the quarterback on the football team has graduated and needs to be replaced. Perhaps students came from forty-nine states the previous year and accepting a certain candidate would mean that the university could claim that it had students from all fifty states. Or perhaps a valued faculty member who is supremely rational in his physics
work threatens to find another job unless his daughter is admitted. In admissions, ideals are continually bumping into reality, and the chosen solution often is not perfectly principled but ultimately is made for pragmatic reasons. Admissions is probably not a good place for idealists who cannot compromise, nor for pragmatists who lack ideals.