THE NEW YORKER

A CRITIC AT LARGE DECEMBER 17, 2001 ISSUE

EXAMINED LIFE

What Stanley H. Kaplan taught us about the S.A.T.

By Malcolm Gladwell

December 10, 2001

Once, in fourth grade, Stanley Kaplan got a B-plus on his report card and was so stunned that he wandered aimlessly around the neighborhood, ashamed to show his mother. This was in Brooklyn, on Avenue K in Flatbush, between the wars. Kaplan's father, Julius, was from Slutsk, in Belorussia, and ran a plumbing and heating business. His mother, Ericka, ninety pounds and four feet eight, was the granddaughter of the chief rabbi of the synagogue of Prague, and Stanley loved to sit next to her on the front porch, immersed in his schoolbooks while his friends were off playing stickball. Stanley Kaplan had Mrs. Holman for fifth grade, and when she quizzed the class on math equations, he would shout out the answers. If other students were having problems, Stanley would take out pencil and paper and pull them aside. He would offer them a dime, sometimes, if they would just sit and listen. In high school, he would take over algebra class, and the other kids, passing him in the hall, would call him Teach. One classmate, Aimee Rubin, was having so much trouble with math that she was in danger of being dropped from the National Honor Society. Kaplan offered to help her, and she scored a ninety-five on her next exam. He tutored a troubled eleven-year-old named Bob Linker, and Bob Linker ended up a successful businessman. In Kaplan's sophomore year at City College, he got a C in biology and was so certain that there had been a mistake that he marched in to see the professor and proved that his true grade, an A, had accidentally been switched with that of another, not quite so studious, Stanley Kaplan. Thereafter, he became Stanley H. Kaplan, and when people asked him what the "H" stood for he would say "Higher scores!" or, with a sly wink, "Preparation!" He graduated Phi Beta Kappa and hung a shingle outside his parents' house on Avenue K—"Stanley H. Kaplan Educational Center"— and started tutoring kids in the basement. In 1946, a high-school junior named Elizabeth, from Coney Island, came to him for help on an exam he was unfamiliar with. It was called the Scholastic Aptitude Test, and from that moment forward the business of getting into college in America was never quite the same.

The S.A.T., at that point, was just beginning to go into widespread use. Unlike existing academic exams, it was intended to measure innate ability—not what a student had learned but what a student was capable of learning—and it stated clearly in the instructions that "cramming or last-minute reviewing" was pointless. Kaplan was puzzled. In Flatbush you always studied for tests. He gave Elizabeth pages of math problems and reading-comprehension drills. He grilled her over and over, doing what the S.A.T. said should not be done. And what happened? On test day, she found the S.A.T. "a piece of cake," and promptly told all her friends, and her friends told their friends, and soon word of Stanley H. Kaplan had spread throughout Brooklyn.

A few years later, Kaplan married Rita Gwirtzman, who had grown up a mile away, and in 1951 they moved to a two-story brick-and-stucco house on Bedford Avenue, a block from his alma mater, James Madison High School. He renovated his basement, dividing it into classrooms. When the basement got too crowded, he rented a podiatrist's office near King's Highway, at the Brighton Beach subway stop. In the nineteen-seventies, he went national, setting up educational programs throughout the country, creating an S.A.T.-preparation industry that soon became crowded with tutoring companies and study manuals. Kaplan has now written a memoir, "Test Pilot" (Simon & Schuster; $19), which has as its subtitle "How I Broke Testing Barriers for Millions of Students and Caused a Sonic Boom in the Business of Education." That actually understates his importance. Stanley Kaplan changed the rules of the game.

The S.A.T. is now seventy-five years old, and it is in trouble. Earlier this year, the University of California—the nation's largest public-university system—stunned the educational world by proposing a move toward a "holistic" admissions system, which would mean abandoning its heavy reliance on standardized-test scores. The school backed up its proposal with a devastating statistical analysis, arguing that the S.A.T. is virtually useless as a tool for making admissions decisions.

The report focussed on what is called predictive validity, a statistical measure of how well a high-school student's performance in any given test or program predicts his or her performance as a college freshman. If you wanted to, for instance, you could calculate the predictive validity of prowess at Scrabble, or the number of books a student reads in his senior year, or, more obviously, high-school grades. What the Educational Testing Service (which creates the S.A.T.) and the College Board (which oversees it) have always argued is that most performance measures are so subjective and unreliable that only by adding aptitude-test scores into the admissions equation can a college be sure it is picking the right students.

This is what the U.C. study disputed. It compared the predictive validity of three numbers: a student's high-school G.P.A., his or her score on the S.A.T. (or, as it is formally known, the S.A.T. I), and his or her score on what is known as the S.A.T. II, which is a so-called achievement test, aimed at gauging mastery of specific areas of the high-school curriculum. Drawing on the transcripts of seventy-eight thousand University of California freshmen from 1996 through 1999, the report found that, over all, the most useful statistic in predicting freshman grades was the S.A.T. II, which explained sixteen per cent of the "variance" (which is another measure of predictive validity). The second most useful was high-school G.P.A., at 15.4 per cent. The S.A.T. was the least useful, at 13.3 per cent. Combining high-school G.P.A. and the S.A.T. II explained 22.2 per cent of the variance in freshman grades. Adding in S.A.T. I scores increased that number by only 0.1 per cent. Nor was the S.A.T. better at what one would have thought was its strong suit: identifying high-potential students from bad schools. In fact, the study found that achievement tests were ten times more useful than the S.A.T. in predicting the success of students from similar backgrounds. "Achievement tests are fairer to students because they measure accomplishment rather than promise," Richard Atkinson, the president of the University of California, told a conference on college admissions last month. "They can be used to improve performance; they are less vulnerable to charges of cultural or socioeconomic bias; and they are more appropriate for schools because they set clear curricular guidelines and clarify what is important for students to learn. Most important, they tell students that a college education is within the reach of anyone with the talent and determination to succeed."

This argument has been made before, of course. The S.A.T. has been under attack, for one reason or another, since its inception. But what is happening now is different. The University of California is one of the largest single customers of the S.A.T. It was the U.C. system's decision, in 1968, to adopt the S.A.T. that affirmed the test's national prominence in the first place. If U.C. defects from the S.A.T., it is not hard to imagine it being followed by a stampede of other colleges. Seventy-five years ago, the S.A.T. was instituted because we were more interested, as a society, in what a student was capable of learning than in what he had already learned. Now, apparently, we have changed our minds—and few people bear more responsibility for that shift than Stanley H. Kaplan.

ADVERTISEMENT

From the moment he set up shop on Avenue K, Stanley Kaplan was a pariah in the educational world. Once, in 1956, he went to a meeting for parents and teachers at a local high school to discuss the upcoming S.A.T., and one of the teachers leading the meeting pointed his finger at Kaplan and shouted, "I refuse to continue until THAT MAN leaves the room." When Kaplan claimed that his students routinely improved their scores by a hundred points or more, he was denounced by the testing establishment as a "quack" and "the cram king" and a "snake oil salesman." At the Educational Testing Service, "it was a cherished assumption that the S.A.T. was uncoachable," Nicholas Lemann writes in his history of the S.A.T., "The Big Test":

> The whole idea of psychometrics was that mental tests are a measurement of a psychical property of the brain, analogous to taking a blood sample. By definition, the test-taker could not affect the result. More particularly, E.T.S.'s main point of pride about the S.A.T. was its extremely high test-retest reliability, one of the best that any standardized test had ever achieved. . . . So confident of the S.A.T.'s reliability was E.T.S. that the basic technique it developed for catching cheaters was simply to compare first and second scores, and to mount an investigation in the case of any very large increase. E.T.S. was sure that substantially increasing one's score could be accomplished only by nefarious means.

But Kaplan wasn't cheating. His great contribution was to prove that the S.A.T. was eminently coachable—that whatever it was that the test was measuring was less like a blood sample than like a heart rate, a vital sign that could be altered through the right exercises. In those days, for instance, the test was a secret. Students walking in to take the S.A.T. were often in a state of terrified ignorance about what to expect. (It wasn't until the early eighties that the E.T.S. was forced to release copies of old test questions to the public.) So Kaplan would have "Thank Goodness It's Over" pizza parties after each S.A.T. As his students talked about the questions they had faced, he and his staff would listen and take notes, trying to get a sense of how better to structure their coaching. "Every night I stayed up past midnight writing new questions and study materials," he writes. "I spent hours trying to understand the design of the test, trying to think like the test makers, anticipating the types of questions my students would face." His notes were typed up the next day, cranked out on a Gestetner machine, hung to dry in the office, then snatched off the line and given to waiting students. If students knew what the S.A.T. was like, he reasoned, they would be more confident. They could skip the instructions and save time. They could learn how to pace themselves. They would guess more intelligently. (For a question with five choices, a right answer is worth one point but a wrong answer results in minus one-quarter of a point—which is why students were always warned that guessing was penalized. In reality, of course, if a student can eliminate even one obviously wrong possibility from the list of choices, guessing becomes an intelligent strategy.) The S.A.T. was a test devised by a particular institution, by a particular kind of person, operating from a particular mind-set. It had an ideology, and Kaplan realized that anyone who understood that ideology would have a tremendous advantage.

Critics of the S.A.T. have long made a kind of parlor game of seeing how many questions on the reading-comprehension section (where a passage is followed by a series of multiple-choice questions about its meaning) can be answered without reading the passage. David Owen, in the anti-S.A.T. account "None of the Above," gives the following example, adapted from an actual S.A.T. exam:

1. The main idea of the passage is that:

A) a constricted view of [this novel] is natural and acceptable

B) a novel should not depict a vanished society

C) a good novel is an intellectual rather than an emotional experience

D) many readers have seen only the comedy [in this novel]

E) [this novel] should be read with sensitivity and an open mind

If you've never seen an S.A.T. before, it might be difficult to guess the right answer. But if, through practice and exposure, you have managed to assimilate the ideology of the S.A.T.—the kind of decent, middlebrow earnestness that permeates the test—it's possible to develop a kind of gut feeling for the right answer, the confidence to predict, in the pressure and rush of examination time, what the S.A.T. is looking for. A is suspiciously postmodern. B is far too dogmatic. C is something that you would never say to an eager, college-bound student. Is it D? Perhaps, but D seems too small a point. It's probably E—and, sure enough, it is.

With that in mind, try this question:

2. The author of [this passage] implies that a work of art is properly judged on the basis of its:

A) universality of human experience truthfully recorded

B) popularity and critical acclaim in its own age

C) openness to varied interpretations, including seemingly contradictory ones

D) avoidance of political and social issues of minor importance

E) continued popularity through different eras and with different societies

Is it any surprise that the answer is A? Bob Schaeffer, the public education director of the anti-test group FairTest, says that when he got a copy of the latest version of the S.A.T. the first thing he did was try the reading comprehension section blind. He got twelve out of thirteen questions right.

The math portion of the S.A.T. is perhaps a better example of how coachable the test can be. Here is another question, cited by Owen, from an old S.A.T.:

In how many different color combinations can 3 balls be painted if each ball is painted one color and there are 3 colors available? (Order is not considered; e.g. red, blue, red is considered the same combination as red, red, blue.)

A) 4 B) 6 C) 9 D) 10 E) 27

This was, Owen points out, the twenty-fifth question in a twenty-five-question math section. S.A.T.s—like virtually all standardized tests—rank their math questions from easiest to hardest. If the hardest questions came first, the theory goes, weaker students would be so intimidated as they began the test that they might throw up their hands in despair. So this is a “hard” question. The second thing to understand about the S.A.T. is that it only really works if good students get the hard questions right and poor students get the hard questions wrong. If anyone can guess or blunder his way into the right answer to a hard question, then the test isn’t doing its job. So this is the second clue: the answer to this question must not be something that an average student might blunder into answering correctly. With these two facts in mind, Owen says, don’t focus on the question. Just look at the numbers: there are three balls and three colors. The average student is most likely to guess by doing one of three things—adding three and three, multiplying three times three, or, if he is feeling more adventurous, multiplying three by three by three. So six, nine, and twenty-seven are out. That leaves four and ten. Now, he says, read the problem. It can’t be four, since anyone can think of more than four combinations. The correct answer must be D, 10.

ADVERTISEMENT

Does being able to answer that question mean that a student has a greater “aptitude” for math? Of course not. It just means that he had a clever teacher. Kaplan once determined that the testmakers were fond of geometric problems involving the Pythagorean theorem. So an entire generation of Kaplan students were taught “boo, boo, boo, square root of two,” to help them remember how the Pythagorean formula applies to an isosceles right triangle. “It was usually not lack of ability,” Kaplan writes, “but poor study habits, inadequate instruction or a combination of the two that jeopardized students’ performance.” The S.A.T. was not an aptitude test at all.

In proving that the S.A.T. was coachable, Stanley Kaplan did something else, which was of even greater importance. He undermined the use of aptitude tests as a means of social engineering. In the years immediately before and after the First World War, for instance, the country’s élite colleges faced what became known as “the Jewish problem.” They were being inundated with the children of Eastern European Jewish immigrants. These students came from the lower middle class and they disrupted the genteel Wasp sensibility that had been so much a part of the Ivy League tradition.They were guilty of “underliving and overworking.” In the words of one writer, they “worked far into each night [and] their lessons next morning were letter perfect.” They were “socially untrained,” one Harvard professor wrote, “and their bodily habits are not good.” But how could a college keep Jews out? Columbia University had a policy that the New York State Regents Examinations—the statewide curriculum-based high-school-graduation examination—could be used as the basis for admission, and the plain truth was that Jews did extraordinarily well on the Regents Exams. One solution was simply to put a quota on the number of Jews, which is what Harvard explored. The other idea, which Columbia followed, was to require applicants to take an aptitude test. According to Herbert Hawkes, the dean of Columbia College during this period, because the typical Jewish student was simply a “grind,” who excelled on the Regents Exams because he worked so hard, a test of innate intelligence would put him back in his place. “We have not eliminated boys because they were Jews and do not propose to do so,” Hawkes wrote in 1918:

> We have honestly attempted to eliminate the lowest grade of applicant and it turns out that a good many of the low grade men are New York City Jews. It is a fact that boys of foreign parentage who have no background in many cases attempt to educate themselves beyond their intelligence. Their accomplishment is over 100% of their ability on account of their tremendous energy and ambition. I do not believe however that a College would do well to admit too many men of low mentality who have ambition but not brains.

Today, Hawkes’s anti-Semitism seems absurd, but he was by no means the last person to look to aptitude tests as a means of separating ambition from brains. The great selling point of the S.A.T. has always been that it promises to reveal whether the high-school senior with a 3.0 G.P.A. is someone who could have done much better if he had been properly educated or someone who is already at the limit of his abilities. We want to know that information because, like Hawkes, we prefer naturals to grinds: we think that people who achieve based on vast reserves of innate ability are somehow more promising and more worthy than those who simply work hard.

VIDEO FROM THE NEW YORKER

What Border Life Looks Like

But is this distinction real? Some years ago, a group headed by the British psychologist John Sloboda conducted a study of musical talent. The group looked at two hundred and fifty-six young musicians, between the ages of ten and sixteen, drawn from élite music academies and public-school music programs alike. They interviewed all the students and their parents and recorded how each student did in England’s national music-examination system, which, the researchers felt, gave them a relatively objective measure of musical ability. “What we found was that the best predictor of where you were on that scale was the number of hours practiced,” Sloboda says. This is, if you think about it, a little hard to believe. We conceive musical ability to be a “talent”—people have an aptitude for music—and so it would make sense that some number of students could excel at the music exam without practicing very much. Yet Sloboda couldn’t find any. The kids who scored the best on the test were, on average, practicing *eight hundred per cent more* than the kids at the bottom. “People have this idea that there are those who learn better than others, can get further on less effort,” Sloboda says. “On average, our data refuted that. Whether you’re a dropout or at the best school, where you end up can be predicted by how much you practice.”

Sloboda found another striking similarity among the “musical” children. They all had parents who were unusually invested in their musical education. It wasn't necessarily the case that the parents were themselves musicians or musically inclined. It was simply that they wanted their children to be that way. "The parents of the high achievers did things that most parents just don't do," he said. "They didn't simply drop their child at the door of the teacher. They went into the practice room. They took notes on what the teacher said, and when they got home they would say, Remember when your teacher said do this and that. There was a huge amount of time and motivational investment by the parents."

Does this mean that there is no such thing as musical talent? Of course not. Most of those hardworking children with pushy parents aren't going to turn out to be Itzhak Perlmans; some will be second violinists in their community orchestra. The point is that when it comes to a relatively well-defined and structured task—like playing an instrument or taking an exam—how hard you work and how supportive your parents are have a lot more to do with success than we ordinarily imagine. Ability cannot be separated from effort. The testmakers never understood that, which is why they thought they could weed out the grinds. But educators increasingly do, and that is why college admissions are now in such upheaval. The Texas state-university system, for example, has, since 1997, automatically admitted any student who places in the top ten per cent of his or her high-school class—regardless of S.A.T. score. Critics of the policy said that it would open the door to students from marginal schools whose S.A.T. scores would normally have been too low for admission to the University of Texas—and that is exactly what happened. But so what? The "top ten percenters," as they are known, may have lower S.A.T. scores, but they get excellent grades. In fact, their college G.P.A.s are the equal of students who scored two hundred to three hundred points higher on the S.A.T. In other words, the determination and hard work that propel someone to the top of his high-school class—even in cases where that high school is impoverished—are more important to succeeding in college (and, for that matter, in life) than whatever abstract quality the S.A.T. purports to measure. The importance of the Texas experience cannot be overstated. Here, at last, is an intelligent alternative to affirmative action, a way to find successful minority students without sacrificing academic performance. But we would never have got this far without Stanley Kaplan—without someone first coming along and puncturing the mystique of the S.A.T. "Acquiring test-taking skills is the same as learning to play the piano or ride a bicycle," Kaplan writes. "It requires practice, practice, practice. Repetition breeds familiarity. Familiarity breeds confidence." In this, as in so many things, the grind *was* the natural.

ADVERTISEMENT

To read Kaplan's memoir is to be struck by what a representative figure he was in the postwar sociological miracle that was Jewish Brooklyn. This is the lower-middle-class, second- and third-generation immigrant world, stretching from Prospect Park to Sheepshead Bay, that ended up peopling the upper reaches of American professional life. Thousands of students from those neighborhoods made their way through Kaplan's classroom in the fifties and sixties, many along what Kaplan calls the "heavily traveled path" from Brooklyn to Cornell, Yale, and the University of Michigan. Kaplan writes of one student who increased his score by three hundred and forty points, and ended up with a Ph.D. and a position as a scientist at Xerox. "Debbie" improved her S.A.T. by five hundred points, got into the University of Chicago, and earned a Ph.D. in clinical psychology. Arthur Levine, the president of Teachers College at Columbia University, raised his S.A.T.s by two hundred and eighty-two points, "making it possible," he writes on the book's jacket, "for me to attend a better university than I ever would have imagined." Charles Schumer, the senior senator from New York, studied while he worked the mimeograph machine in Kaplan's office, and ended up with close to a perfect sixteen hundred.

These students faced a system designed to thwart the hard worker, and what did they do? They got together with their pushy parents and outworked it. Kaplan says that he knew a "strapping athlete who became physically ill before taking the S.A.T. because his mother was so demanding." There was the mother who called him to say, "Mr. Kaplan, I think I'm going to commit suicide. My son made only a 1000 on the S.A.T." "One mother wanted her straight-A son to have an extra edge, so she brought him to my basement for years for private tutoring in basic subjects," Kaplan recalls. "He was extremely bright and today is one of the country's most successful ophthalmologists." Another student was "so nervous that his mother accompanied him to class armed with a supply of terry-cloth towels. She stood outside the classroom and when he emerged from our class sessions dripping in sweat, she wiped him dry and then nudged him back into the classroom." Then, of course, there was the formidable four-foot-eight figure of Ericka Kaplan, granddaughter of the chief rabbi of the synagogue of Prague. "My mother was a perfectionist whether she was keeping the company books or setting the dinner table," Kaplan writes, still in her thrall today. "She was my best cheerleader, the reason I performed so well, and I constantly strove to please her." What chance did even the most artfully constructed S.A.T. have against the mothers of Brooklyn?

Stanley Kaplan graduated No. 2 in his class at City College, and won the school's Award for Excellence in Natural Sciences. He wanted to be a doctor, and he applied to five medical schools, confident that he would be accepted. To his shock, he was rejected by every single one. Medical schools did not take public colleges like City College seriously. More important, in the forties there was a limit to how many Jews they were willing to accept. "The term 'meritocracy'—or success based on merit rather than heritage, wealth, or social status—wasn't even coined yet," Kaplan writes, "and the methods of selecting students based on talent, not privilege, were still evolving."

That's why Stanley Kaplan was always pained by those who thought that what went on in his basement was somehow subversive. He loved the S.A.T. He thought that the test gave people like him the best chance of overcoming discrimination. As he saw it, he was simply giving the middle-class students of Brooklyn the same shot at a bright future that their counterparts in the private schools of Manhattan had. In 1983, after years of hostility, the College Board invited him to speak at its annual convention. It was one of the highlights of Kaplan's life. "Never, in my wildest dreams," he began, "did I ever think I'd be speaking to you here today."

The truth is, however, that Stanley Kaplan was wrong. What he did in his basement *was* subversive. The S.A.T. was designed as an abstract intellectual tool. It never occurred to its makers that aptitude was a social matter: that what people were capable of was affected by what they knew, and what they knew was affected by what they were taught, and what they were taught was affected by the industry of their teachers and parents. And if what the S.A.T. was measuring, in no small part, was the industry of teachers and parents, then what did it mean? Stanley Kaplan may have loved the S.A.T. But when he stood up and recited "boo, boo, boo, square root of two," he killed it. ♦

Published in the print edition of the December 17, 2001, issue.

Malcolm Gladwell has been a staff writer for the The New Yorker since 1996.