NEW YORK TIMES BESTSELLER

"This beautifully written, defiantly hopeful book points the way to a better future for American teachers and the children they teach."
—Paul Tough, best-selling author of *How Children Succeed*

Elizabeth Green

BUILDING

BETTER TEACHER

HOW TEACHING WORKS

(and How to Teach It to Everyone)

with a new afterword

1

FOUNDING FATHERS

By 1948, when he landed his first academic job at the University of Illinois, Nate Gage had already helped the army select and train radar observers during World War II; worked with the College Board to develop a new tool—the Scholastic Aptitude Test; and coauthored a definitive textbook: *A Practical Introduction to Evaluation and Measurement*. The second son of Jewish immigrants from Poland, he'd made his way from hanging wallpaper with his father to the top of his chosen field, educational psychology. At Illinois, he joined the prestigious new Bureau of Educational Research. But the breakthrough that became Nate's most important finding happened in the classroom.

Nate was serious, but also passionate and sweet. At conferences, he would transfix his colleagues with barroom storytelling late into the night. And yet, in the classroom, that chemistry somehow failed to materialize. He simply could not keep the students' attention. It was not unusual for one or more of them to fall asleep in the middle of his lectures. "He just didn't have that certain *something*," says one of his students, David Berliner. For

all his success—the multiple publications in prestigious journals; the glittering title, *professor of education*—the data all pointed to one disturbing conclusion: Nate was a terrible teacher.

Distraught, Nate turned to the academic literature. Surely some of his colleagues in educational psychology had cracked the mysteries of teaching. That was when he made his second discovery: the research on teaching didn't exist. At least, the *findings* didn't. Instead of conclusions, researchers had developed a bundle of idiosyncratic hypotheses, focused mostly on teachers' personality traits. Were good teachers warmer? more enthusiastic? more organized? more interested in their subject? Maybe better teachers had similar degrees of bohemianism, emotional sensitivity, and sociability. Perhaps subpar teachers displayed radicalism, or even "worrying suspiciousness." Other studies cast their searches even more broadly, investigating traits from age and experience to eye color, clothing style, and strength of grip.

None of the studies found anything conclusive. A researcher would publish a discovery, only to have another produce exactly opposite findings. The few conclusions that could be squeezed out of the research tended to be vague and unhelpful. One set of studies suggested that good teachers should be "friendly, cheerful, sympathetic, and morally virtuous rather than cruel, depressed, unsympathetic, and morally depraved." Another study concluded that the best teachers had a characteristic called, unhelpfully, "teaching skill."

Summarizing the research in 1953, Nate wrote:

> The simple fact of the matter is that, after 40 years of research on teacher effectiveness during which a vast number of studies have been carried out, one can point to few outcomes that a superintendent of schools can safely employ in hiring a teacher or granting him tenure, that an

> agency can employ in certifying teachers, or that a teacher-education faculty can employ in planning or improving teacher-education programs.

The irony was bruising. The country, at that point, had dozens of university programs devoted to recruiting, training, and vouching for America's future teachers—education schools, they were called. Yet somehow all those ed schools' professors had managed to learn nothing about teaching. And that was the professors who paid the topic any attention at all. The most prestigious among them—the elite education researchers like Nate—ignored teaching altogether.

You couldn't help but wonder. How had this happened? How had an entire field come to neglect the work at its heart?

One answer was that they did it on purpose. The tradition began with the first education professors, who taught the new education courses with undisguised reluctance. "Educational psychology?" the philosopher William James was said to have quipped. "I think there are about six weeks of it." James became the grandfather of the discipline. His student, Edward Thorndike, another foundational figure, entered the field only because he had to. After he finished graduate school in psychology in 1898, the best job offer he could find was not in psychology but in pedagogy, at the Women's College at Western Reserve University in Cleveland.

"The bane of my life is the practice school they stuck me with," he wrote in a letter to a friend soon after starting the job. Later, when he moved to Columbia University's Teachers College, he spent his first year visiting schools, but he quickly abandoned the mission, calling the trips a "bore." When asked what he would do if faced with a certain superintendent's real-world dilemma, he scoffed. "Do? Why, I'd resign!"

Instead of addressing educational problems, Thorndike took psychological ones and grafted them onto schools. He applied to human students the general laws of learning that he derived from his experiments with monkeys, dogs, and cats ("Never will you get a better psychological subject than a hungry cat," he wrote). Meanwhile, he aided the proliferation of new measurement techniques, assessing everything from intelligence to memory. But he did not study teachers.

Even John Dewey, who advocated a "science of education," wound up retreating to his original discipline, philosophy. All around him, educational researchers had followed Thorndike and abandoned the study of real schools. Discouraged, Dewey set his work in education aside.

Nate Gage, too, never intended to study education. What he really wanted to be was a psychologist. But after graduating from the University of Minnesota magna cum laude, a star student of the young B. F. Skinner, he was rejected by all ten graduate programs he applied to. "From the universities' point of view it would be pointless to take him into a graduate programme in psychology and waste resources training him, since he was Jewish," explained Minnesota's dean of psychology, Richard Elliott. Graduate programs were judged by their success at placing professors, and universities did not hire Jews. The only program that made him an offer was one he had not applied to—a new program in educational psychology at Purdue, where the young director recruited his students by scouring psych departments' reject lists.

Another reason early education professors ignored teaching was that they found it uninteresting. Learning to teach composition did not require a method, but rather a "clear head, an enduring conscience, an elastic enthusiasm, and uncommon commonsense," the English professor LeBaron Russell Briggs

insisted. "There is no such thing as a science of Pedagogy," Josiah Royce wrote in the lead article of the inaugural issue of the journal *Educational Review*, published in 1891. "As for a 'philosophy of education' in any other sense," Royce added, "the lord deliver us therefrom."

Yet the subject had to be offered; simple economics demanded it. In 1890, total enrollment in US elementary and secondary schools stood at just under thirteen million. By 1920, the number was more than twenty million. In the same period, the ranks of school teachers grew by nearly four hundred thousand. Another twenty-one thousand people served as administrators. By the time Nate arrived at the University of Illinois, in 1948, the number of teachers alone was nearing one million. For a university, the calculation was clear: training teachers made financial sense whether there was something to teach them or not.

The grim history might have led another man to surrender. If William James hadn't been able to develop a science of teaching, what could honestly be expected of Nate Gage from New Jersey? Anyway, as Thorndike had proved, it was perfectly possible to make a respectable career in education research without touching the teaching problem at all. But where others might have seen a dead end, Nate saw possibility. After all, in science, the most important discoveries were born not from answers, but from puzzles. And, studying the early work on teaching, he had glimpsed a common and, he suspected, fatal flaw.

None of the traits the first researchers investigated—eye color? strength of grip?—had come from the classroom. They had looked into hundreds of variables but ignored "the primary data of the teaching process." That choice, too, belied the pattern of science's greatest discoveries. Johannes Kepler, Dmitri Mendeleev, Gregor Mendel—all began by scrutinizing phenomena close up and

only *then* came up with theories to explain them. Like "Kepler in examining the orbits of planets, Mendelyeev [*sic*] in poring over the properties of the elements, or Mendel in raising his peas," Nate decided, education researchers would only unlock the mysteries of the American classroom by venturing inside of it.

Nate set out to construct a true science of teaching. He called his method the "process-product" paradigm. By comparing the process (teaching) to its product (learning), researchers could conclude which teaching acts were effective and which were not. The ambition was not unlike John Dewey's imagined science of "what the gifted teacher does intuitively." The only difference was that, while Dewey favored learning about teaching in the messy cauldron of a real school, Nate preferred formal experimentation. A successful process-product study, in his view, needed to approximate the natural classroom habitat while also controlling for extraneous variables.

In one experiment, Nate focused on explanation, the slice of teaching that, in his opinion, formed "the essence of instruction." He and his grad students recruited real teachers to teach real students, but under certain parameters. One was that the teachers could speak and use the chalkboard, but they could not invite discussion, solicit questions, or even ask students to take notes. ("For some teachers this restriction may require a difficult departure from their customary teaching style," the instructions read. "We hope that you will bear with us.") Another restriction was the content; each lesson corresponded to a preselected article from the *Atlantic Monthly* magazine. The researchers gave students a comprehension test at the end of the lesson to find out which teachers had explained it best.

Nate's students videotaped each lesson and catalogued the teachers' behaviors. One graduate student, Barak Rosenshine, had a list of twenty-seven qualities to watch for, ranging from

the average length of words spoken (perhaps brevity was key?) to the frequency of "reference to pupils' interests" to the number of gestures ("movement of the arms, head, or trunk") and paces (walking from one place to another).

Another group wrote computer programs to analyze what the teachers had said. One compared the transcripts against a "vagueness dictionary" written specially for the occasion (qualifying words included *almost*, *maybe*, *generally*, and *most*). In one lecture that scored as highly vague, for instance, a teacher began by describing an author's name, which he said was "not too important." He went on:

> I will put his name up on the board anyway. It is really not very important at all. MIHAJOV [*sic*]—that is the way you pronounce that word, Uh Mihajlov wrote those articles. And someone, he has done something that is fine someone very similar had done and there was another author whose name, uh, uh, let us just remember there is another author. That one has spelling problems too. Two authors, two authors. One we know is Mihajlov, the other one wrote earlier in nineteen sixty-two. Both of them complained about conditions, especially in Russia. And this one was in prison because he wrote a book about conversations with Stalin and, I do not know if you have ever heard of the book . . .

The final step compared the recorded teaching behaviors (process) and students' comprehension scores (products). As one might expect, the students of vagueness offenders had significantly lower comprehension. Rosenshine's method yielded other strong correlations. A high number of gestures, it turned out, helped improve comprehension; so did a high level of right-to-left movement. The research might not have been quite what Dewey

imagined, but it was certainly unlike anything Nate Gage's contemporaries had seen.

Process-product research caught on quickly. In 1957, sharing an elevator with a colleague at the American Educational Research Association's annual conference, Nate joked that if the elevator crashed, then all of that year's research on teaching would go down with it. That year, he and the colleague were the only two giving papers on the topic. By the spring of 1963, Nate's book collecting the available research on teaching had converted a new generation of researchers into the fold. Officially called *The Handbook of Research on Teaching* (and unofficially known as "The Gage Handbook"), the volume sold 30,000 copies. One chapter, outlining how to design experiments to study teaching, generated such demand that the publisher, Rand McNally, released it separately in pamphlet form in 1966. By 1974, the pamphlet had sold 130,000 copies.

Perhaps most important, Nate became, if not the most engaging teacher, certainly a beloved one. Graduate students devoted themselves to him, and even the American Federation of Teachers, a union representing practitioners across the country, caught on. "They called him the Sage Gage," says Lovely Billups, a union official at the time, who worked with Nate to convert his findings into usable lessons for teachers.

So when, in 1971, a pair of young staffers at the new National Institute of Education was charged with funding the next generation of research on teaching, they went straight to teaching's "pooh-bah," according to one of them, Garry McDaniels. Soon, Nate was taking a leave from his university—by then he was at Stanford—to help them launch the new round of funding with a conference suggesting new directions for research.

There was just one twist. Created by the contrarian new president, Richard Nixon, NIE was charged not just with supporting

existing research, but with transforming it. "My assignment," says McDaniels, "was to change the field." Wittingly or not, Nate helped him do it. The draft conference agenda he circulated for feedback went to all his colleagues back at Stanford, including the man who would eventually inherit Nate's "pooh-bah" crown—a young professor visiting from Michigan State named Lee Shulman.

"Garbage," Lee Shulman said when Richard Snow, another Stanford professor, asked him what he thought of Gage's draft—the one he was circulating about the conference planning the future of research on teaching. "Same old bullshit."

Dick Snow was aghast. "Why?"

"It's nothing but a kind of testimony to the past," Lee said. "Doesn't Nate realize that behaviorism is on life support?"

It was true. Nate's process-product approach depended on a school of psychology that was falling increasingly out of fashion. Nate was a behaviorist by default and also by generation; B. F. Skinner, his old professor, had been behaviorism's seminal figure. The founder of educational psychology, Thorndike, was another lifelong adherent. Nate's rise correlated with behaviorism's most prominent period.

The behaviorists held that the only scientific way to study humans was to study their directly observable features—their behaviors and the actions ("stimuli") that triggered them. But the new generation of psychologists began to point out that by focusing on stimuli and their responses, behaviorists were ignoring the mind.

In Thorndike's model, the human mind was just an extension of the animal one. Learning meant responding to repeated rewards or punishments. If rewarded for one behavior enough times, the subject learned to keep doing it. If punished, he or she learned to stop.

But while this pattern might describe some forms of human learning, critics argued that behaviorism could never explain them all—especially not the kind of learning that went beyond simple actions (will I get food when I press my cat paw on this pedal?) to more complicated concepts (when is it useful to calculate an indefinite integral?). To explain how people learned higher-level concepts, the critics held, psychology had to reckon with cognition.

Lee, who'd begun not as a psychologist but as a philosopher, had never liked behaviorism. It rejected as unscientific the questions that he found most fascinating—questions about the mind. Early on, that opinion was unpopular. But by the time of Lee's year at Stanford, in 1973, critics—known as "cognitivists"—had broken the behaviorist stranglehold on their field. The cognitive revolution spread from one area of psychology to the next, turning attention from behavior to the working of the mind.

Lee figured the shift should apply to research on teaching too. The whole point of process-product research, Nate Gage's great contribution, was to study teaching by studying teachers' behaviors. But what about their minds?

"Why don't you write to Nate?" Snow told Lee.

"Come on, Nate *personifies* process-product research!" Lee said.

But Snow was insistent. Nate was a serious scholar. He'd listen. "So I wrote him a two-page memo," Lee says. "Probably wrote it on a typewriter, Selectric typewriter, and I made—I politely critiqued what he was doing and said, 'You don't even have one group looking at the relevance of cognitive work for the study of teaching, and my guess is that's the future of research on teaching.'"

Lee was mostly just riffing. "I mean, I wasn't really in the field at that point. I was teaching future teachers . . . But research on

teaching wasn't my area." So when the phone rang a few days later and Nate asked him to use the memo as the basis for leading one of the ten panels at his conference in Washington, Lee was unprepared. He didn't think. He just said yes.

Lee Shulman's area of expertise was doctors. He'd begun studying them in 1968, at Michigan State, as an outgrowth of an idea that first struck him in graduate school.

Besides education, what Lee had always found fascinating was thinking. The technical term was *epistemology*, the occupation of thinking about thinking. Like his idol John Dewey, Lee focused on higher kinds of thoughts, the mental operations that take place when a person moves from impression to question to understanding. "The pedestrian," wrote Dewey, "*feels* the cold; he *thinks* of clouds and a coming shower."

The psychology of thinking wasn't just fascinating; it also seemed painfully relevant to education. By understanding complex thought—the process of making knowledge—researchers would not just study schools; they would help improve them. And Lee had an idea for how to study thinking in a way that could make a real difference. Other early cognitive psychologists presented subjects with problems to solve, puzzles to answer, but Lee knew that, in real life, problems didn't come prepackaged. "A problem well put is half solved," John Dewey wrote. "Without a problem, there is blind groping in the dark." To get a true grasp on how knowledge was made, Lee intended to study the blind groping in action. He only had to find the right research subjects—people for whom problem solving was part of the natural habitat.

The idea of studying doctors arrived a few years into Lee's time at Michigan State, when a man walked into his office and introduced himself as the dean of the university's new medical

school. "I understand you study complex problem solving," he said to Lee. "Well," he continued, "I think that's what medicine's all about, and we physicians don't begin to understand how that really works. Would you be willing to take 50 percent of your appointment and join the medical school faculty and do research on medical problem solving?"

Doctors. Of course! "It was such an epiphany," Lee says. Doctors solved problems all day. It was the heart of their work. Joseph Bell, Sir Arthur Conan Doyle's medical school professor and a surgeon with legendary capacities of deduction, had inspired Sherlock Holmes, the greatest professional problem solver in (fictional) history.

Lee said yes, and it was a perfect fit. Observing doctors at work with his colleague and childhood friend Arthur Elstein, he overturned the conventional wisdom about medical problem solving—and, ultimately, helped improve medical education in the process. Lee and Arthur designed simulations to approximate the circumstances of daily diagnosis and asked doctors to discuss their thought processes. Students played the patients. A lab room became the doctor's office, staged like a regular exam room except for the two huge video cameras mounted on the ceiling. Three real cases provided the basis for the actors' improvisation, and Lee and Arthur concocted a "data bank" with all the blood levels and X-ray results a physician might possibly request. As the doctors worked, researchers stood behind a one-way mirror watching them "think aloud," sharing the mental considerations that usually remain private.

On the first day, Lee, Arthur, and their colleagues got a preview of what they would find. Watching their first physician, a chief of medicine, the researchers expected events to proceed as all the medical textbooks recommended. First the doctor would interview the "patient." Then he would start ordering tests. Only

later, after reviewing the results, would the doctor start outlining possible diagnoses.

But the work-up had barely begun when the chief of medicine turned to the researchers to announce his first diagnosis. What was going on? At first the team figured the chief of medicine must be a maverick, an outlier who followed his instincts. But as more doctors came into the lab, each one proceeded in a similar manner, suggesting two, three, even four possible diagnoses before even taking the patient's blood pressure. The maverick wasn't a maverick at all. The majority of doctors worked this way, exactly the opposite of the meticulous decision tree that textbooks advised.

But the method seemed to work. When Lee, Arthur, and their team ran their data, they found that doctors who made their first diagnostic guess earlier in the appointment got the answer right just as often as those who waited. If anything, it looked like the more guesses were made early on, the more likely the physician was to reach an accurate diagnosis. So much for moving "from symptom to sign to syndrome to disease," as one textbook prescribed. With one modest study, Lee and his team had discovered that medical decision making was far more complex than the textbooks portrayed.

Lee thought he could take the research even further. At Stanford, that was what he planned to do—extend the problem-solving findings, fleshing out their implications for education. And that, ultimately, is what he did. He just didn't realize quite what form the transformation would take.

After the NIE conference, writing up a report based on his panel discussion, Lee's first move was to borrow from his own work, crossing out the word *physician* and writing *teacher* instead. The clinical act of medical diagnosis became the clinical act of teach-

ing; the questions about which lab tests to run became questions about how to group the students, arrange the classroom, and select a textbook. Where Nate had thought of teachers as collections of behaviors, Lee borrowed from the medical project and called them "information processors."

Lee had no expectations for his foray into the study of teaching. Cognitive scientists had started out by studying doctors, chess masters, and investors because thinking was an obvious prerequisite of their job. How much information was processed by people who spent their days telling small children, "One, two, three, eyes on me!"?

But studying teachers by studying their thinking turned out to be surprisingly generative. The process-product findings that Nate Gage championed might have been statistically significant, but they often seemed to contradict each other. It was important that every child stay "on task," but calling on students at random—the best way to keep them focused—was not always the best path to getting a good discussion going. Similarly, after asking a question, the most successful teachers waited a few extra seconds before accepting an answer. But successful teachers also tended to be the most brisk, spending the smallest number of minutes between topics. Pulling a single, clear answer out of the process-product research was like trying to distill laws from the Bible. One passage offered perfect clarity, but the next said the complete opposite.

Lee, who had spent his grade-school years at a yeshiva, met the task as perhaps no other psychologist could. "Think about the tradition of commentaries on the Talmud—this enormously long historical tradition of interpretation in which you never get to a settled conclusion," says Gary Sykes, who worked with Lee at Stanford. "It's brilliant intellectual work with a text. Lee took as the text intellectual life in classrooms. And from there, all was commentary and interpretation."

Take the problem of timing. How could it possibly be beneficial both to be fast, moving quickly from task to task, and also to be slow, pausing beyond the bounds of comfort before calling on a student to answer your question? Lee explored this teaching problem by examining what he called "the anatomy of a turn." The process-product researchers had described the visible elements of turns: teacher asks a question, time elapses, student answers. But to really understand the turn, you had to look at it from the teacher's perspective.

Building on others' research, especially Mary Budd Rowe's study of "wait time"—the pause between posing a question and selecting an answer—Lee pointed out the logic in the apparent paradox. For a teacher, each second spent waiting for an answer held both promise and danger. On one hand, the longer she waited, the more time the students would have to think. This was good. On the other hand, the sooner she broke the silence with the correct answer, the lower was her risk of exposing the class to a useless diversion. This was also good. The wisdom and peril of pausing were both true, and if you thought that didn't make sense, well, that was true too. Wait times, Lee concluded, were "blessings dipped in acid."

The question for teachers, as for doctors, was not, What is the best behavior? It was, How do I decide which of many behaviors to deploy for the case at hand? It was a problem of diagnosis. Teachers had to locate their pupils' pathologies, determine a best intervention, and act.

With doctors, diagnosis and treatment had clear beginnings, middles, and ends. With teachers, the questions kept coming. Since the pathologies—that is, everything the child didn't know—were not physical but mental, how could teachers diagnose them? How could they understand what a child had failed to learn? And if they did manage to teach successfully, how could they confirm it?

There was also the problem of scale. "The teacher," Lee realized, "is confronted not with a single patient, but with a classroom filled with 25 to 35 youngsters." Even if a teacher could locate pathologies and somehow do it for all her students, how did she manage to deploy the correct interventions, all at once, to the entire group? "The only time a physician could possibly encounter a situation of comparable complexity," Lee concluded, "would be in the emergency room of a hospital during or after a natural disaster." Studying teachers, he realized, was just as important as thinking about doctors; in fact, "it is far more germane."

The National Institute of Education conference came and went quickly. Lee moderated his panel; submitted his summary report, advocating the usefulness of studying teachers' decisions; and soon he was back at Michigan State, working with doctors. He might have forgotten about the trip altogether, had NIE not sent him a call for proposals to build a new research and development center to study teacher thinking and decision making.

Lee knew that his proposal would be a long shot. The likely list of applicants included Stanford University and his colleague Nate Gage. And since writing his famous handbook, Nate had made Stanford into the country's leading source of research on teaching. MSU, by comparison, was a "cow college," better known for training teachers than for studying them.

But Michigan State won. Among the losers were several of the universities that had been pulling government grants for behaviorist education research for years, Gage among them. "Nate lost his grant too," says Garry McDaniels. "In the old days they always gave him the grant. But the work that he had done had been going on for so long that I was convinced that it had reached its end."

Lee is fond of quoting a line from the psychologist Jerome

Bruner about narrative. One of the cognitive revolution's leaders and an early scholar of teaching, Bruner wrote that narrative is fundamentally composed of "the vicissitudes of intention." A protagonist sets out to do one thing, but along the way something unpredictable happens, and he decides to do another thing instead.

Lee set out to study thinking. By understanding minds, he thought, he could help improve education, the work of shaping them. The thing that happened along his way—the call from Nate Gage—led him to change not his intention, but his method. Doctors had provided a neat keyhole into the mind, but it turned out that another group of professionals offered a bay window. Teachers not only had to think; they had to think about other people's thinking. They were an army of everyday epistemologists, forced to consider what it meant to know something and then reproduce that transformation in their students. Teaching was more than story time on the rug. It was the highest form of knowing.

At a university, traditionally the highest degree holders are called *master* or *doctor*. "Both words," Lee discovered, "have the same definition; they mean 'teacher.'" What was the best way to show you really understood a subject, if not to teach it? And what was the best way to use research to improve education, if not to study teaching?

Without realizing what he was doing, Lee had stumbled on Dewey's lost project. Teaching was indeed the science of all sciences, the art of all arts, as Dewey's predecessor Francis Parker had put it. And now, thanks to Nate Gage's nudge, Nixon's investment, and his own lifelong obsession, Lee was going to pick up on the work Dewey and Parker had never finished.

Lee had written in the NIE panel's concluding report that "gifted practitioners are capable of performances which our best theories are not yet capable of explaining, much less generating

or predicting." Future research on teaching, then, should explore the talents of the best teachers—the "wisdom of practice," he called it. All he needed to do was find the great practitioners.

Lee Shulman was no teacher, but he became one of two seminal figures in modern thinking about policies for improving the quality of teaching. The other figure was not a teacher either. And whereas Lee focused on education after spending time inside of it (or at least inside of a school of education), Eric Hanushek came to it wholly from the outside. He would go on to have as much of an influence on education as Lee did, if not more. But he never worked at an ed school.

Hanushek became fascinated with schools in the summer of 1966. He was nearing the end of graduate school in economics at MIT, still lacking a dissertation topic, when he stumbled on a remarkable story in the newspaper:

> **Washington, D.C.**—The Johnson administration Thursday was accused of ignoring results of a Federal investigation of inequality in city school systems because of political implications . . .
>
> After a survey of 600,000 school children and 60,000 teachers, the report concluded that pupils from poor families left school "with greater deficiencies" than when they entered . . .
>
> "This report means that all our education plans—increasing spending per pupil, more and better libraries and books, education devices—won't solve the crisis in our schools," said [Connecticut congressman Abraham] Ribicoff.

Given the redistributive goals of Lyndon Johnson's "Great Society" programs, it made sense that the administration would want to cover up the report. If the study, by a Johns Hopkins sociologist named James Coleman, was right, then one of the most expensive educational interventions in history had failed. According to Coleman, giving schools additional services—including more per-pupil spending, the supposed antidote to underachievement—did not help poor and African-American students overcome the challenges of their environments.

Hanushek couldn't quite believe it. "If in fact schools don't make much difference," he thought, "why are we continually pumping more and more money into schools to try and improve them?" There had to be something else going on, a lurking variable masking the money's impact. But what could it be? Running through Coleman's data, a massive set drawing from 645,000 students and more than three thousand schools across ninety-three different variables, Hanushek found no mistakes of consequence. Nor did a working group convened at Harvard to vet the research. Coleman's conclusions largely held up.

But what about other data? Hanushek pulled together a data set from a school district in California—much smaller in scope than Coleman's national sample, but with two advantages. First, instead of capturing a snapshot of one year in students' lives, the California data followed students longitudinally. Second, the data broke down students not just by school, but by the teachers they'd had. The extra detail enabled Hanushek to get more specific than Coleman; he could go beyond whether schools made a difference and determine whether *individual teachers* had an effect as well.

The effect of teachers was no simple thing to measure, even with the better data. Countless factors undoubtedly influenced

students' performance in schools, from genetic disposition to the size of their parents' vocabularies. How could Hanushek discern the teacher's influence amid all these other variables?

The education literature offered no advice, but another area of economics did: the study of industrial production. Like teachers, factories receive certain raw products (steel, coal, plastics) and then put their own unique spin on the business of transforming them into, say, a Chevy Camaro. To measure the productivity of the manufacturing process, economists had to extract the value offered by raw products from the value provided by the plant assembling them. They did this by looking for patterns. What was the value of the raw products before they came to the factory, and how much did that value rise or fall after the manufacturer had its way with them?

Applying the idea to education, Hanushek could control for the effects of nonteacher variables, from home background to past performance, by searching for deviations. "If you follow an individual kid and you see him on some learning path, and then one year, all of a sudden, he learns a lot more than in another year, or all of a sudden he learns a lot less," Hanushek explains, "that gives you a hint that maybe it's something specific about the teachers, or something specific in that year. Then, if you see that all of the kids in the one class have this jump or this fall in performance, then you start to believe that it's something in that class."

The California data was full of such jumps. If Hanushek's method was right, teachers did make a difference, and the difference was big. Later he managed to put a number to the effect. Students assigned to the best teachers, he calculated, progressed by the equivalent of a whole grade level more than students assigned to the worst, as measured by test scores.

By Hanushek's method, teachers could do what the Coleman Report suggested schools could not: they could offset the disadvantage of poverty.

Perhaps Hanushek's most influential finding stemmed from his comparison of teachers' "effectiveness" (the educational equivalent of productivity) to other characteristics, especially salary. The amounts paid to teachers were based on how many years of experience they had and on how many degrees they had earned; a master's won a salary bump, and every additional graduate class won another. Yet the salary inputs seemed to have no bearing on the output: teacher effectiveness. Experience did matter, but only up to a certain point. In productivity terms, there was no difference between a teacher who'd been teaching for three years and a teacher who'd been teaching for thirteen.

Writing his final dissertation, eventually published as a book called *Education and Race*, Hanushek turned the observation into a suggestion—one that would reverberate for years into the future. If school districts stopped rewarding graduate study and experience, then they could redirect their investments into something more efficient: "Teacher Accountability," he called it.

Accountability would draw on the statistical method that Hanushek had adapted from studies of factories. "This procedure," he explained, "allows the ranking of teachers on the basis of teaching ability." Ranking and then rewarding teachers according to their effectiveness might create some problems—including, he suggested, "problems arising from attempts to 'teach the tests.'" (Held accountable for how their students performed on tests, teachers might emphasize the exam material to the exclusion of other, equally important lessons.) But, Hanushek wrote, "while we may not be at a point now where we trust standardized tests to hold up under concerted attempts to

foil them, conceptually the problem appears soluble." A decade later, Hanushek gave his method a name: "value-added."*

In the same book, *Education and Race*, Hanushek made one more intriguing point. Like Coleman, he had only examined educational investments and their effects—comparing, in economic parlance, education's "inputs" to its "outputs." He had not studied education's vast middle. "The black box of the production process," he called it. That is, classroom teaching and learning. He had, in other words, followed the trail that Nate Gage and Lee Shulman were blazing, looking not just at schools but at teachers, yet he had followed them only so far. He looked at teachers' effects, but not at their work—at teachers, but not at teaching.

Hanushek made the observation as an aside, but the decision to overlook teaching's "black box" would prove just as influential as his "value-added" innovation. By studying teaching, Lee Shulman and his colleagues were about to explode many common ideas about how it worked, including the myth of the natural-born teacher. Hanushek, meanwhile, ignored teaching and, as a result, ignored how teaching worked. He could read his value-added research and draw the simplest conclusion, the one that matched what everyone already believed: some teachers were bad, most were fine, and a few were wonderful—as if they'd been born that way.

* The other researcher with a claim to having invented value-added calculations of teacher effectiveness is the statistician William Sanders, who developed the Tennessee Value-Added Assessment System (TVASS).