Bridging the Gap Between Research and Practice:

WHAT'S GOOD, WHAT'S BAD, AND HOW CAN ONE BE SURE?

Mr. Davis, a professor, advises readers of this article who are practitioners not to trust everything he and his colleagues have to say about schools. Then he explains why and offers helpful tips that will allow teachers and administrators to make their own judgments about what they can profitably apply from educational research.

BY STEPHEN H. DAVIS

> We have to live today by what truth we can get today and be ready tomorrow to call it falsehood.
> — *William James*

I TEACH at a large research university and deeply appreciate the opportunities I have to help shape the careers of aspiring school leaders and, in some ways, the institution of public education itself. As expected of professors in universities like mine, I have written quite a bit and conducted several research studies that I hope have made a reasonable contribution to my field. But I often wonder how much of what I've written has actually made a difference in the lives of school practitioners. For that matter, I wonder how much of what is thought about and produced by scholars and researchers actually affects the way administrators and teachers behave in schools. And most important of all, I wonder how practitioners learn to judge the quality of research and determine what research to pay attention to and what to ignore. These questions are crucial, especially now that the mandates of No Child Left Behind and the standards and accountability movement are pressuring American public schools to use research-based programs.

Feedback from the field suggests that a gap between research and practice persists while bridges between them remain tenuous and unsteady. It appears that comparatively little of what is written and thought about by scholars and policy makers actually has any appreciable impact on classrooms or drives durable systemwide reform efforts. In their acclaimed book on school reform, *Tinkering Toward Utopia*, Larry Cuban and David Tyack trace the long and often confounding history of reform efforts in American public schools.[1] They argue that, despite decades of reform initiatives and millions of dollars spent in the pursuit of educational innovation, the fundamental tenets of effective educational practice have changed very little.

Of course, not all reform efforts have been research-based, and not all good research is lost in the trickle-down journey between the halls of academe and Ms. Doe's third-grade classroom. But enough of value is lost to

STEPHEN H. DAVIS is an associate professor in the School of Education, Stanford University, Stanford, Calif.

raise suspicions about the relevance of the work of researchers and the vitality of the relationship between researchers and public school practitioners.

For decades, public schools in America have been awash in waves of reform that have emerged from a slippery mélange of empirical research studies, politically formulated mandates, and locally derived best-practice initiatives. Most have been well intentioned, many have been misguided, and some have come to schools without careful regard for hard evidence or with distorted claims of causality. Unfortunately, some of what has made it into Ms. Doe's classroom may not represent the best research, nor has it necessarily been applied with fidelity to the scientifically supportable findings from which it was derived. A close look at how research findings relating to such topics as heterogeneous classrooms, mainstreaming of special-needs students, social promotion, bilingual education, and instructional methods are actually applied in public school classrooms reveals numerous variations based on local policies and politics, management philosophies, school culture, student characteristics, levels of teachers' skill, and available resources.

But today, now that public schools have entered an era of high-stakes accountability and standards-based instruction in which decisions about educational programs and activities are expected to be closely aligned with empirical research and evidentiary data, it can be argued that scholars and practitioners must redouble their efforts to bridge the gap between theory and practice.[2] Unfortunately, conversations about the reform du jour usually overlook any questions of how to tell the difference between good research and not-so-good research or between the attributes of research in the abstract and research as it may apply to a particular school district, school, or classroom. Perhaps most important, even research of the highest quality may never make its way into public school classrooms simply because the pipeline through which important academic discoveries travel to schools and classrooms is inconsistent, inchoate, and often tainted by the political process used to craft education policy.

For example, critics maintain that the death of bilingual instruction in California public schools in the 1990s was more the product of ill-informed public sentiment and partisan politics than of a serious and balanced investigation of empirical research on how children who are not native speakers of English learn a new language. The fact is that bilingual education came under heavy fire because of its costs, the expected growth of new financial burdens, the paucity of basic research in the field, the lack of qualified teachers, and its real or supposed lack of success.[3]

As a former superintendent, principal, personnel director, and high school teacher, I can't count the times that I've heard fellow administrators and teachers begin a conversation about the qualities of a particular educational program or instructional method with the phrase "The research says." It is always uttered as if the mere declaration, in and of itself, contains sufficient gravity to forestall further dissent or discussion. The problem, of course, is that most school administrators and teachers are not sophisticated consumers of research. Nor are many college professors, for that matter. Practitioners, especially, take far too much for granted and often exaggerate the importance of mediocre research, attempt to generalize qualitative research findings, and frequently misapply research findings derived from one sample population to different populations. Most people (scholars included) also tend to pay greater attention to research studies that confirm their deeply held beliefs and assumptions than to studies that challenge their sensibilities.[4]

So how can practitioners learn to become better consumers of scholarly research? Short of conscripting teachers and administrators into doctoral-level courses in research methods and short of reassigning all education professors to a team-teaching stint with Ms. Doe, I offer some thoughts and suggestions that may help to inform practitioners about the factors that influence the quality and evaluation of educational research. In the process, I hope to reduce the distance between the worlds of practitioners and academics.

THE ARROGANCE OF ACADEME

To begin, school leaders should never completely trust what people like me have to say about schools. I'm serious. Take what we professors have to say with a grain of salt. I know, most of you are thinking that I must have had a particularly bad day at the office. But, truth be told, I'm one of "them." By them, I mean professors and researchers who ply their trade from the rarefied air of the ivory tower. What we have to tell you about schools simply can't be fully trusted. Here's why:

1. Many of us have simply lost touch with the day-to-day complexity of human interactions in schools. As a result, too many of us begin to think about schooling and the behaviors of people who work and study in schools as abstractions, as plots on a regression line, or as categorical representations of patterned activity. Too many of us have buffered ourselves and our work from the subtleties, nuances, and untidiness of human behavior in schools. Academic myopia can cloud our ability to appreciate on a visceral level what life in schools is really like and, more important, the extent to which grand ideas generated from our research actually stick to the ribs of school reform efforts.

Recently, I spent some time on the campus of a large suburban high school, leading a scheduled state accreditation review process. Over the years, I've volunteered for "accreditation duty" numerous times and have visited many schools. Each experience has enriched my understanding of how schools in the U.S. work. Yet each experience also underscores my growing sense of disconnection from the real people who populate schools, from the turbulent rhythm of school life, and from the logic of school operations (i.e., the relationship between events, situations, and objects and the inevitable consequences of their interaction). I spent the better part of 19 years teaching and leading in public schools, but I am incrementally losing touch. And I know I can't be alone in this. Imagine a scholar who has never worked in the school system and whose only contacts have been narrowly tailored to the needs of his or her particular research interests. Academic pedigrees and good intentions notwithstanding, would you trust what this person has to say about schools?

In my nationally acclaimed school of education, I am the *only* professor who has been a principal or superintendent of a large, comprehensive public school or school district.

Here's a sobering fact: in my nationally acclaimed school of education, I am the *only* professor who has been a principal or superintendent of a large, comprehensive public school or school district. Only a modest number of faculty members have actually taught in public schools, and, for those who have, many years have passed since their public school teaching days. Yet several colleagues are deeply involved in researching, teaching, and consulting on topics related to the leadership, governance, and redesign of schools. For many of my colleagues, their understanding about life in schools is based on their exposure to the literature, their own particular strand of research, episodic campus visits, or the anecdotal reports of those who work in schools. A rough parallel might be a physician whose knowledge of the inner workings of the human body was derived primarily from examining the pictures in Gray's *Anatomy* and supplemented by occasional visits to the operating room viewing theater. A great deal is lost when "book learnin'" forms the primary source of a person's expertise.

My colleagues here and in similar universities across the country are unquestionably accomplished scholars with brilliant minds, and the best of them often generate new thoughts and ideas that ought to be carefully considered by practitioners. But should practitioners completely trust what they have to say about schools?

2. Academics often write for the wrong audience. As a group, academics gain notoriety, respect, and career advancement not from their ability to reach out to the "grunts on the front lines" but from their ability to impress their peers. Young professors don't acquire tenure by writing for trade journals, and associate professors who aspire to become full professors don't make the leap by publishing books based on literature reviews or philosophical musings about best practices in schools. Rather, we write for one another, while adhering to the strictures of academic scholarship. In the process, we often fail to adequately capture or even acknowledge the nuanced qualities of schooling. Granted, every so often, seeds of wisdom escape this incestuous process to take root in the terra firma of schools, but not necessarily through the intentional efforts of their originators. As a result, practitioners are justifiably suspicious about the relevance of scholarly research. A review of several widely read education journals is more likely to reveal articles written by practitioners for practitioners than articles written by academics for practitioners.

A big piece of the communication gap is in the language used by scholars. The language we speak is different from that spoken by common folk. For example, a school administrator or teacher not schooled in the lexicon of statistics (and most aren't) would probably find the following statement from a recently published article in *Educational Administration Quarterly* nearly incomprehensible: "Multilevel modeling provides a framework in which researchers can place explanatory variables at their correct level of the data hierarchy." Here's another brain buster from the same journal: "Organizations do not appear ex nihilo but instead emerge within an existing sociopolitical context. How this context interacts with organizational factors to initiate normative structure is poorly understood." Poorly understood? Why am I not surprised?

The language gap between researchers and practitioners is underscored by the following quote from Carolyn Riehl and her colleagues: "Research should be accessible to its community. Educational administration professors are often frustrated by the fact that many students find the language of research off-putting. But in fact, much of the research is written in language and terms that make it inaccessible to many of the people it is intended to reach."[5]

Conscientious scholars can and should write to be understood. Education experts such as Diane Ravitch, Michael Fullan, Michael Kirst, Larry Cuban, and Linda Darling-Ham-

mond are among the most prominent and respected scholars in their fields. Yet they have mastered the craft of communicating complex ideas in ways that can be understood by practitioners without compromising the sophistication of their scholarship.

A group discussion at a recent faculty meeting shows the low regard accorded to scholars who write for practitioners. The work of a promising candidate for a tenure-track faculty position was under close scrutiny at the meeting. The candidate had amassed an impressive array of published articles in refereed journals reporting the results of her extensive body of empirical research. However, one transgression nearly derailed her candidacy: she was in the finishing stages of writing a book intended for the parents of public school children, not for scholars. The fact that the book was written well and filled a void did little to assuage the prickly indignation expressed by some senior faculty members.

3. Academics frequently use "hit and run" tactics. We are harvesters of data and rarely stay very long in one place. Although some scholars do indeed conduct long-term and penetrating studies of their subjects and environments, far too many of us rely on ex post facto analyses of behaviors or events that occurred at a single time. As a result, what we often present are snapshots of events that fail to capture the patterns, routines, trends, or rhythms that represent organizational life or individual behavior.

Scholars often fail to consider the synergistic properties of organizational activities in schools or take account of the unanticipated effects of subtle relationships. We tend to focus instead on relatively narrow dimensions of human activity and thinly sliced representations of organizational or leadership behavior. Unfortunately, researchers who operate this way can engender feelings of mistrust and suspicion about their motives, a situation that impedes free and forthright dialogue with practitioners. Seasoned practitioners tacitly understand the holistic properties of schools, and they are loath to give much credibility to scholars with narrowly framed research agendas.

4. Academics can be an arrogant bunch. Swathed in egos plumped up by the untouchable status that comes with lifetime tenure, academics love to pontificate to the masses with unshakable certainty of the righteousness of their beliefs. Their messages are often interpreted as "We're basically smarter than most of you, and therefore we know more than you do." Do they really? The corridor of history is littered with discarded bodies of work that were at the time of their invention considered to be on the cutting edge of scholarly research. Practitioners should listen carefully to what we have to say, but always assume a critical stance. They need to ask themselves whether an idea will work in their school or district, given the unique needs and characteristics of the students, teachers, and community members. They need to decide whether the idea aligns with common sense, and they need to remember that, just because it's new, it doesn't mean it's necessarily better.

5. Not all research is good research. You'd be surprised at the quantity of poorly supported information presented in the guise of empirical research that has made its way to the front lines of public schools. The literature is rife with half-truths, popular myths, contradictions, poorly designed studies, misinterpreted findings, and conclusions soaked in the personal biases and deeply held assumptions of researchers.[6]

I am reminded of the body of research relating to the field of organizational leadership. Thirty years ago, the literature commonly portrayed successful leaders as charismatic and heroic figures. They were men (most often) who possessed the requisite ambition, vision, drive, and know-how to steer their organizations toward desirable ends. Today, prominent authors like Thomas Sergiovanni and Linda Lambert portray a very different model of leadership that depicts successful leaders as humble servants and moral stewards who are transformational, inspirational, and skilled in the ability to facilitate collaboration, common vision, a sense of community, and shared values.[7] What proof do we have about the effectiveness of either model? Do more recent models of leadership result in better student learning? How can practitioners make sense of the contradictory "evidence" about effective leadership?

Perhaps a better example of how research myths can permeate the collective mentality of educators is the emergent theory that large comprehensive high schools are no longer the most effective vehicle for educating teens. The dominant argument is that large schools are impersonal and overly bureaucratic and too easily alienate students — particularly those who are at risk or come from underrepresented ethnic groups. To "fix" this problem, school districts across the country have begun to convert large high schools into smaller learning communities. Such efforts are under way in the inner cities, in the suburbs, and even in some rural communities — despite the fact that the idea was originally intended as a way to advance the academic achievement and graduation rates of low-income urban youths. In an unexpected move, the Bill & Melinda Gates Foundation, benefactor to more than 1,600 small-school conversion programs, recently toned down its energetic support for the concept, maintaining that perhaps a more important strategy would be to focus on improving the instructional practices of teachers.

On its face, the logic of creating smaller schools is compelling: greater faculty empowerment in school governance

and decision making, increased personalization for students, less pressure on students to conform to large-group norms, less exposure to disruptive behaviors, improved feelings of identification with the school, better response time in meeting individual needs, increased flexibility of program and personnel policies, more frequent opportunities for faculty collaboration, and better monitoring of individual student progress.

Attractive in theory, but does it really work? Moreover, is the model necessarily superior to more traditionally structured comprehensive schools? In some environments, it certainly is; in others, not necessarily. Should models of conversion to small schools become a universal template for high school reform? Consider the potential tradeoffs: fewer elective and specialty courses, reduction in the breadth and scope of staff expertise, fewer opportunities for extracurricular and athletic participation, increased intra- and extra-organizational competition for resources, limited access to specialized instructional facilities like science labs, and an increase in the likelihood of policy drift within school districts (i.e., departures from common values, goals, and quality indicators). Moreover, do the assumptions that underlie the logic of small schools necessarily apply evenly to all communities, all students, and all socioeconomic groups? Do teachers actually teach differently in small schools than they do in larger schools? Are these differences large enough to foster better student learning? Is the curriculum offered to students in small schools appreciably better than the curriculum commonly offered in larger schools? Does increased personalization always result in better student academic achievement? Will the "treatment" effects of small school conversions justify the costs in all school settings (e.g., for upper-income suburban students as opposed to lower-income urban students)? What will the consequences be for those students who are already doing well in large comprehensive school settings?

These and many other questions remain largely unanswered and should give pause to practitioners who might otherwise interpret the recent outpouring of favorable literature on the subject as proof positive of the approach's universal efficacy. According to a recent study by Barbara Schneider, Adam Wysse, and Vanessa Keesler of Michigan State University, there's consensus in the field that it's not simply going smaller, it's what you do and how you teach in those smaller settings that matter most.[8] A recent report published by the National Association of Secondary School Principals concludes by stating, "Implementing small learning communities will not, in and of itself, increase student achievement. It may help to do so, but the studies do not provide conclusive evidence on this point."[9]

Some research, of course, is simply crummy, regardless of the topic under scrutiny. For example, it is not uncommon for authors to overstate the importance of their findings. In other cases, authors have been known to highlight findings that agree with their deeply held values and assumptions while downplaying the importance of contradictory findings. It is also not unheard of for authors to use weak or inappropriate statistical measures, to use self-designed surveys and questionnaires that are unreliable and lack validity, or to generalize findings derived from nonrandomized samples or from qualitative data.

I am reminded of the highly controversial (and by many accounts slippery) findings of Richard Herrnstein and Charles Murray, described in *The Bell Curve*, in which the authors maintained that intelligence is both inheritable and differentiated by race. Or the book *The End of Racism*, in which Dinesh D'Souza claimed that racism is no longer an important factor in American life and that the government must therefore cease to legislate issues on a racial basis.[10]

How can practitioners uninitiated in the finer points of research design distinguish between the good, the bad, and the just plain ugly? One way to start is to read research published in reputable journals (usually refereed). For example, in the field of educational leadership, practitioners could rely on such journals as *Educational Administration Quarterly*, the *Journal of School Leadership*, and the *Journal of Educational Administration*. Other examples of reputable journals in education include *Educational Researcher*, *Harvard Educational Review*, and the *American Journal of Education*. Another way is for practitioners to become proficient by learning the fundamentals of research design from credible experts like John Creswell or Michael Patton. And yet a

third way is to read more than one article on a particular subject and then compare and contrast methods, findings, and conclusions. Look for common threads, discrepant information, and citations of supporting work throughout the text. However, regardless of how good one becomes at this, always keep in mind that even the best research can be wrong for a particular school district, school, or classroom.

QUANTITATIVE OR QUALITATIVE: WHICH IS BETTER?

Thirty years ago, reputable research in the field of education was normally quantitative in design. The predominant research method was to distribute surveys to randomly selected samples of targeted subjects and to conduct statistical comparisons of numerical datasets. In fact, most scholars in education considered qualitative research methods to be "soft," lacking in methodological rigor, and of limited value with regard to the study of large populations. It was widely — and not altogether unreasonably — believed that qualitative research was susceptible to researcher bias and the generation of unsubstantiated and highly subjective data based largely on self-reported perceptions and experiences. The question of how one can reliably extrapolate useful information from a qualitative study and confidently apply the findings to large numbers of schools and children continues to be at the center of the debate.[11]

Many contemporary scholars maintain that good qualitative research has equaled, if not exceeded, quantitative research in status, relevance, and methodological rigor.

However, much has changed. In many universities today, the popularity and frequency of qualitative research in education overshadow the traditional emphasis on quantitative studies. Many contemporary scholars maintain that good qualitative research has equaled, if not exceeded, quantitative research in status, relevance, and methodological rigor. In essence, over the past quarter century, education scholars came to the realization that most surveys and statistical analyses failed to capture the fine-grained qualities of schooling. Yes, well-designed quantitative studies allowed for generalizing findings to larger populations, but legitimate concerns arose regarding the ability of such research to effectively capture the nuances of human interactions and program effects, differences in environmental contexts, and depth of understanding. Today, the focus is on finding patterns of behavior, documenting through thick descriptions, and seeking deeper meaning rather than higher levels of statistical significance.

Despite the rise in the popularity of qualitative research, a pronounced shift has recently occurred in national policy. In 2001 the No Child Left Behind Act refocused the national agenda of educational researchers on studies that can be classified as "scientifically based." To determine whether a study is scientifically based, NCLB offered the following standards:

- The research involves the application of rigorous, systematic, and objective procedures to obtain reliable and valid knowledge.
- The research employs empirical methods that draw on observation or experiment.
- The research employs rigorous data analyses adequate to test stated hypotheses and justify general conclusions.
- The research problems and data are evaluated using experimental or quasi-experimental methods.

The question of whether qualitative research can be considered scientifically based according to these standards is the center of a vigorous debate among scholars and policy makers.[12] And the jury is still out. So that brings us back to the question of what is better, quantitative or qualitative research? Perhaps the question shouldn't rest on the notion of what is better, but what is most relevant to practitioners in the field.

Here's my take on the matter. Since practitioners deal mostly with the day-to-day intimacies and idiosyncratic behaviors of children and colleagues, they are best served by research findings that can help them build technical competence and strategies for managing other people. Statistical analyses and numerical data can fail to provide the fine-grained insights needed to complete a full portrait of how schools work and how learning occurs in nuanced environments. Therefore, well-devised qualitative research (e.g., ethnographic studies, case studies, phenomenologies, grounded theories, and other descriptive narratives) is most likely to resonate with practitioners — and especially with practitioners not well trained in the fundamentals of statistical research (i.e., the vast majority).

Once again, however, I must caution that, regardless of the research design or methods used, the results should always be reviewed with a critical eye.

PRACTITIONERS ARE PART OF THE PROBLEM

Of course, practitioners also contribute to the persistent gap between research and its application in practice.

The items that follow are several errors in thinking about research, truth, and knowledge that are commonly made by practitioners.

1. The seductive power of silver bullets and gurus. Besieged by the unremitting turbulence of daily activities in schools, practitioners favor technical solutions and "quick fixes" and have little patience (or time) for more abstract or nuanced research about schools and students. They want to know what works and how to go about fixing things. The problem is that there are no silver bullets. What works in one setting or with a particular group of people will inevitably play out differently (subtly or significantly) in different settings or with different people. Practitioners are wise to examine scientifically based research, but always with the understanding that, in the education business, "one size rarely fits all."

Practitioners also love gurus. Gurus offer us visions, wisdom, and certainty. Gurus exist in many forms. They can be thought leaders, authors, consultants, professors, politicians, or practitioners. They come wrapped in a veil of magic and often possess a personal magnetism that draws people into their sphere. In general, gurus influence others through the power of their personalities and their ideas.

Thirty years ago, I attended a seminar conducted by Werner Erhard (born John Paul Rosenberg), the spiritual leader of the New Age self-help group EST. It was held at the Cow Palace in Daly City, California. Over ten thousand people sat transfixed for several hours while Erhard spoke extemporaneously about existentialism, self-empowerment, and the meaning of life. Although I recall very little about the content of his talk, I vividly remember his charisma. Through the power of his mesmerizing personality, Erhard virtually ate up the stage. For all I cared, he could have been speaking about the finer points of Etruscan pottery. For thousands of devotees, he was a guru exemplar. Whether his ideas were right or wrong, I don't know. But he certainly had an impassioned point of view and the ability to convince others of the truth of his beliefs. He was later convicted of tax evasion and suspected of spousal abuse — but that's a different story.

Let me share another story to illustrate how the power of a scientifically based idea can cause practitioners in search of silver bullets and enamored by gurus to misapply the idea in practice. At about the same time, in the mid-1970s, UCLA professor Madeline Hunter championed what was to become a wildly popular model for planning and assessing teaching. Referred to as the Seven Step Lesson Plan, Hunter's "direct instruction" model (which she assembled from several research studies) quickly became the silver bullet that administrators and teacher trainers across the nation craved.

Unfortunately, too many school districts and administrators used the model as a rigid template for judging the quality of teaching. By her own admission, Hunter recognized the model's limits and repeatedly cautioned practitioners not to apply it too rigidly. After all, not every lesson necessarily required the application of all seven steps. Nevertheless, teacher evaluation protocols, rating forms, and staff development programs that framed teaching exclusively in terms of these seven steps popped up in districts across the country.[13]

As a young high school administrator in a district that used Hunter's model religiously, I fervently applied the framework with each teacher whom I was responsible for evaluating. A highly popular veteran teacher, Richard Loftus, was my very first victim. Loftus was an economics teacher who loved to lecture. In fact, that's just about all he ever did. Nevertheless, he did it exceptionally well, and the students relished it. Over the years, many of Loftus' former students who went on to college reported that his course more than adequately prepared them for college-level economics. Moreover, on the economics Advanced Placement exam, most of Loftus' students outperformed students across the state. In hindsight, either he had an incredible run of brilliant students, or he was doing something right in the classroom.

During my first observation of Loftus in action, I sat in the back of the classroom, armed with a Seven Step Lesson Plan checklist. At the end of the lesson, he had failed to employ even one of the seven steps. However, his lecture was lively, creative, and informative. Students were taking notes and were highly attentive. During the post-lesson conference, I informed Loftus that his rating for the lesson was unsatisfactory and that he needed to improve in all seven areas. Without uttering a word in his defense, he abruptly stormed out of my office and slammed the door. He rarely spoke with me again. Although Loftus wasn't much of a "scientist" in his approach to teaching, he was an exceptional "artist," a quality that was completely overlooked by the rigid application of the district's evaluation protocol.

Gurus aren't all bad. In fact, some are the real deal. They're often smart, experienced, and wise. But who are the modern-day gurus — Tom Peters, Bill Gates, Maya Angelou, Hillary Clinton, Al Gore, Margaret Wheatley, the Dalai Lama? The answer is: all of them and none of them. It depends on one's point of view, the context, and situational factors. Not one, however, possesses a "silver bullet."

One last comment about gurus: gurus are gurus not always because they challenge us to think differently, but often because what they offer resonates with our existing

values and beliefs. So we must pick and choose among them and sort out the relevant from the irrelevant, the meaningful from the nonsensical, and we must always maintain an underlying sense of skepticism about what we (and they) really know.[14]

2. Exaggerated attributions of causality and misconceptions about chance. Practitioners are prone to overestimate the effects of particular reforms and underestimate the probability of randomness. According to Daniel Kahneman and Amos Tversky, "People often make extreme predictions on the basis of information whose reliability and predictive validity are known to be low."[15] Promising new programs and practices that generate improvements in student learning and performance are often credited with "causing" the desired outcomes. However, educating children can't be reduced to a set of universal truths, skills, or strategies. Education is as much art as it is science.[16]

No reform, no method, no pedagogical approach can completely control for the inevitable and infinite variations of human behavior and their effects on learning. As a result, it is almost impossible to provide irrefutable proof that a particular educational approach acts as the sole causal agent in changed behavior. One must also consider how other factors, such as peer influences, teacher dispositions and skills, physical and emotional status of the child, family influences, the Hawthorne Effect, and so on, may contribute to observed or measured changes. According to a nationally distributed school reform research briefing, "It is important to have high ambitions and reasonable expectations about the results that reforms can produce. Careful evaluations of reform efforts seldom find large and dramatic effects."[17]

3. Presumed associations. Practitioners often assume that, because an empirically supported principle of human behavior or learning or an innovative program or practice seems to work in one setting or context, it will apply just as well in different settings. But is this a safe assumption? For example, will a reading program that works well for a certain group of inner-city minority students elicit comparable levels of growth from minority students in middle-income suburban settings? Will a successful approach to developing English-language skills with immigrant Asian children from developed countries work equally well with immigrant Hispanic children from developing countries? Is the widely held principle that homogeneous grouping of students is bad for learning (and most certainly bad education policy) true in all situations? Will children in homogeneous groups always fail to learn as effectively or as much as children in heterogeneous groups? Will zero tolerance discipline policies reduce campus violence with equal effect in lower-income inner-city schools and middle-class suburban schools? Do the strategies used to encourage school involvement with highly educated parents in white-collar professions work equally well with less-educated blue-collar parents? Does the elimination of social promotion practices in schools stimulate low-achieving students to perform better?

The answer to these and many other questions like them is, "Sometimes yes, and sometimes no." However, one thing is certain: just because an empirically supported practice works well in one context doesn't mean that it will work equally well in other contexts (similar or otherwise). Neither does it mean that the positive effects of a research-based reform will be equally sustainable over the long term from one setting to the next.

4. Wishful thinking. Under the press for sustained improvement in student achievement, educators are generally eager to see new educational interventions succeed. They are far too busy with the day-to-day tasks of managing schools and classrooms to waste time on unproven or weak strategies. So when they do invest themselves in promoting change, it's not uncommon for them to be optimistic that their efforts will pay off.

Take, for example, the tragic case of a charismatic superintendent from a large suburban school district in Northern California. In an effort to improve academic achievement and reduce the dropout rates among poor minority students, he devised a highly innovative and expensive magnet school plan. At the time (the late 1980s), magnet schools had gained considerable popularity across the nation's urban school districts. They were seen as a way to distribute students more evenly across socioeconomic boundaries, pique students' intrinsic motivation, and provide intra-district choice. So eager were the superintendent and the school board to implement the plan that they based several critical decisions about resource allocation on their expectations of supplemental funding from both private and public sources. Tragically, the funding sources never materialized, and the school district went into state receivership, after posting a budget deficit of more than $28 million. The superintendent lost his job, but he was able to repeat this scurrilous feat in two more school districts before his eventual retirement.[18]

The point is that wishful thinking clouds good judgment. When we want something bad enough, it's easy to throw caution to the wind while ignoring disconfirming evidence and common sense.

5. Generalizing from nonrandom and small samples. Jaime Escalante was, in the minds of many educators, a miracle worker. A passionate and creative high school math teacher born in Bolivia, Escalante gained notoriety in the 1980s for his legendary ability to take low-income, under-

achieving minority students at Garfield High School in Los Angeles and transform them into Advanced Placement calculus savants through his magical approach to teaching. He was the subject of the popular movie *Stand and Deliver*, which idealized his exploits and understated the years of work it took to cultivate the math skills of a select group of carefully groomed students. Escalante became a hero to legions of administrators and teachers. Yet, despite countless efforts by inner-city teachers across the country to replicate his instructional success, few have matched or sustained Escalante's track record of success.[19] Why is that?

For one thing, Escalante's methods were applied to very small samples of students. He didn't teach in 10 schools, or a hundred, or a thousand. He taught in one school in Los Angeles and one in Sacramento. And, over the course of his career, not all students succeeded on the Advanced Placement exams. Then, too, Escalante was Escalante. The qualities and behaviors that make a successful teacher are far more complex than the instructional methods used in the classroom. Granted, we all learned a great deal from Escalante about passion, focused instruction, and the application of math concepts to real-world problems. But we also know that the sample of students whom he taught over the years was not necessarily representative of all students in all schools and certainly not large enough to confidently generalize the qualities of his teaching to all teachers who work in high-minority schools in the nation's cities.

The moral of the story is that what works in Ms. Doe's classroom may or may not work in yours. There is no way to reliably predict success. However, what works in the classrooms of a thousand randomly selected Ms. Does is far more likely to work in yours as well.

6. Generalizing from perceptions and self-reported data. Much of the qualitative and quantitative research today depicts self-reported experiences, judgments, and perceptions of reality. Understanding what, how, and why particular individuals think or behave the way that they do by asking them can be quite interesting and useful. Well-written narrative accounts of self-reported data can be fascinating, very persuasive, and authoritative in tone. Such accounts, however, cannot and should not be generalized to large groups.

There are three particularly pernicious aspects of self-reported information: fallible memories, biased mental models, and narrow frames of reference. Let's start with memory. Memories are imperfect. They are distorted by time, lack of use, new stimuli, and emotions. Vivid as a memory might be, it is never objective and always represents an individual's subjective account of reality and truth.[20] Granted, memories are often all a researcher has to go on, but at best they represent approximations of past events.

Self-reported data are deeply influenced by an individual's mental models. In plain language, a mental model represents a person's world view, assumptions, beliefs, and core values. As first explained by psychologist Kenneth Craik in the early 1940s, the mind constructs "small-scale models" of reality that it uses to anticipate events. Mental models can be constructed from perception, imagination, or the comprehension of discourse. Like memories, mental models are unique to individuals and never fully represent objective truth. The way we think and feel about the world shapes our interpretation of events past, present, and future.

Finally, an individual's ability to frame situations and problems from multiple perspectives bears upon the credibility of that individual's interpretation of events. Narrow-minded people see the world through a singular lens and often fail to empathize with others' perspectives or feelings. They also fail to consider alternative explanations of the truth or reality.[21] Some critics of the war in Iraq contend that the current U.S. Administration failed to consider or anticipate the social and political dynamics of the post-Saddam Iraq through multiple frames of reference. As a result, the depth and breadth of the insurgency has greatly exceeded prewar expectations.[22]

To correct for the distortions of memory, mental models, and framing, competent researchers look for corroborating evidence. They evaluate self-reported information in light of the narrative explanations of others, confirming or disconfirming documentation, and observation. Without such careful craftsmanship, self-reported data should always be treated with caution and never extrapolated to explain the behaviors, perceptions, values, or beliefs of large groups.

CREDIBLE EVIDENCE: WHAT TO ATTEND TO AND WHAT TO IGNORE

I'd like to conclude by revisiting the question of what practitioners should attend to and what they should ignore. First, practitioners must consider empirical research as encyclopedic rather than plenary. In education, even the most credible research is subject to differing interpretations and rarely depicts the final word or an indisputable truth. Good research is a road map and rarely a destination.

Second, useful evidence can come in many forms. Obviously, a well-designed research study provides the most credible form of evidence about what works in schools. But whether the research is qualitative or quantitative in form is less important than its relevance to problems of practice. And the most important issue for practitioners is not the broad relevance of research to the field, but the relevance of research findings to particular contexts and circumstances. For example, research about developing

charter schools as vehicles for improving academic achievement among low-income urban minority youths probably won't provide much relevant information to those who want to develop charter schools in middle- and upper-income suburban environments. Likewise, attempts to apply empirically supported strategies for teaching reading to underachieving children from low socioeconomic family backgrounds are likely to be less effective in stimulating the reading skills of underachieving children from higher socioeconomic backgrounds. Stated simply, context matters.

Third, don't dismiss the usefulness of anecdotal evidence. Firsthand accounts of what works and what doesn't are important sources of information. At the very least, such accounts can (and should) provoke inquiry into the research literature. Corroborating the anecdotal evidence with reports from several expert practitioners and from the research literature, of course, is more credible than simply relying on individual anecdotal accounts. However, the intriguing potential of anecdotal evidence is that it can stimulate experimentation and innovation. Since silver bullets are extremely rare in the education business, practitioners must rely on modest experiments and incremental "wins." They must understand that making progress in the education of children is rarely linear and more often recursive, episodic, and even idiosyncratic. Practitioners must develop a "nose" for possibilities, imaginative strategies, and potential pathways that may lead to improved educational practice. The findings of published research alone probably won't get you to the promised land.

Finally, trust your gut. If an empirically tested strategy or program doesn't feel right for your school or district, it probably isn't right. Intuition can be a useful barometer for judging when to slow down and search for additional information. Let common sense be your guide. A recent study of 92 school principals revealed that the principals frequently used their intuition when making important decisions. Most important, the study also revealed that, even though intuitively derived decisions didn't always turn out well, the principals always regretted it when they failed to follow their gut.[23]

---

1. Larry Cuban and David Tyack, *Tinkering Toward Utopia: A Century of Public School Reform* (Cambridge, Mass.: Harvard University Press, 1998).

2. Frances Lawrenz, Douglas Huffman, and Bethann Lavoie, "Implementing and Sustaining Standards-Based Curriculum Reform," *NASSP Bulletin*, June 2005, pp. 2-17.

3. James Crawford, Stephen Krashen, and Haeyoung Kim, "Anti-Bilingual Initiatives: Confusing in Any Language," Hispanic Link News Service, 29 March 1998.

4. Jeffrey Pfeffer and Robert I. Sutton, *Hard Facts, Dangerous Half-Truths, and Total Nonsense: Profiting from Evidence-Based Management* (Cambridge, Mass.: Harvard Business School Press, 2006).

5. Carolyn Riehl et al., "Reconceptualizing Research and Scholarship in Educational Administration: Learning to Know, Knowing to Do, Doing to Learn," *Educational Administration Quarterly*, vol. 36, 2000, p. 405.

6. Pfeffer and Sutton, op. cit.

7. Thomas J. Sergiovanni, *Moral Leadership: Getting to the Heart of School Improvement* (Hoboken, N.J.: Wiley, 1996); and Linda Lambert, *Leadership Capacity for Lasting School Improvement* (Alexandria, Va.: Association for Supervision and Curriculum Development, 2003).

8. Deborah Viadero, "Study Questions Push for Smaller High Schools," *Education Week*, 6 June 2006, pp. 12-13.

9. Janet Quint, "Research-Based Lessons for High School Reform: Findings from Three Models," *Principal's Research Review*, vol. 1, no. 3, 2006, pp. 1-8.

10. Richard Herrnstein and Charles Murray, *The Bell Curve: Intelligence and Class Structure in American Life* (New York: Simon & Schuster, 1994); and Dinesh D'Souza, *The End of Racism: Principles for a Multiracial Society* (New York: Simon & Schuster, 1996).

11. Gavriel Solomon, "Transcending the Qualitative-Quantitative Debate: The Analytic and Systemic Approaches to Educational Research," *Educational Researcher*, August/September 1991, pp. 10-18.

12. Margaret Eisenhart and Lisa Towne, "Contestation and Change in National Policy on 'Scientifically Based' Education Research," *Educational Researcher*, October 2003, pp. 31-38.

13. Patricia Wolfe, "What the 'Seven-Step Lesson Plan' Isn't," *Educational Leadership*, February 1987, pp. 70-71.

14. Pfeffer and Sutton, op. cit.

15. Daniel Kahneman and Amos Tversky, "Intuitive Prediction: Biases and Corrective Procedures," in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982), p. 417.

16. N. L. Gage, *The Scientific Basis of the Art of Teaching* (New York: Teachers College Press, 1978).

17. Quint, p. 7.

18. Lisa Snell, "Scandals Prevalent in Public Schools" (commentary), Reason Foundation, 2004, available at www.reason.org/commentaries/snell_20040601b.shtml.

19. Jay Mathews, *Escalante: The Best Teacher in America* (New York: Holt, 1989).

20. Antonio R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (1994; rpt. New York: Penguin, 2005).

21. Lee G. Bolman and Terrence E. Deal, *Reframing Organizations: Artistry, Choice, and Leadership*, 3rd ed. (New York: Wiley, 2003).

22. Larry J. Diamond, *Squandered Victory: The American Occupation and Bungled Effort to Bring Democracy to Iraq* (New York: Holt, 2005).

23. Stephen H. Davis and Patricia B. Davis, *The Intuitive Dimensions of Administrative Decision Making* (Lanham, Md.: Scarecrow Education, 2003).