Measured Approach or Magical Elixir?

How to Tell Good Science from Bad

ILLUSTRATIONS BY JAMES YANG

By Daniel T. Willingham

Bernhard Dohrmann is a businessman and entrepreneur of wide-ranging interests. Unfortunately, he has also had his share of legal problems. In 1975, he was convicted of securities fraud for selling railroad cars that did not exist. In 1982, he was charged by the Federal Trade Commission with misrepresenting the prices of investment diamonds. The case was settled out of court, with Dohrmann's company returning $6.7 million to investors. In 1991, Dohrmann was charged by the U.S. attorney's office with 16 counts of criminal contempt; it seems he lied about his company's sales figures when selling bonds to investors. He was sentenced to prison for this crime in November 1995.[1]

With such a history of legal problems, what's a troubled businessman to do? Why, go into the educational software business, of course!

Dohrmann started a company called Life Success Academy that marketed (and continues to market) Super Teaching. Super Teaching consists of a system that projects images to three screens; the central screen shows whatever images a teacher typically uses in a lesson plan. The flanking screens show "seemingly random" images of nature, or real-time footage of the teacher or the students. This practice is said to be consistent with "whole brain learning."[2] Systems initially were to sell for $160,000 *per classroom*;[3] the current price is down to $29,500.[4]

Although Super Teaching had been around since at least 2002, things started to look really promising for the Life Success Academy in December 2007, when the company signed an agreement with the University of Alabama in Huntsville. The university would help test and refine the Super Teaching method and would in

---

Daniel T. Willingham is a professor of cognitive psychology at the University of Virginia and the author of numerous articles, including his regular "Ask the Cognitive Scientist" column for American Educator. *To read more of his work on education, go to www.danielwillingham.com. This article is excerpted from his new book,* When Can You Trust the Experts? How to Tell Good Science from Bad in Education, **

return share in profits from future sales. In early October 2008, the university unveiled Super Teaching with a ribbon-cutting ceremony. The president of the university attended, but the honor of cutting the ribbon went—not inappropriately—to Tony Robbins, motivational speaker and late-night infomercial pitchman.[5]

A year and a half later, the University of Alabama in Huntsville dissolved its relationship with Dohrmann and the Life Success Academy.[6] Things had heated up six months earlier. A blog that covers Alabama politics had posted a lengthy summary of Dohrmann's criminal past, provocatively headlined "Why Is UAH Involved with 'a Very Dangerous Con Man'?"[7] A month later, the university's student newspaper published an article titled "Learning at the Speed of Con."[8]

This may be an extreme example, but it's hardly news that an educational reform idea attracted serious attention despite the fact that there was no evidence supporting it. If that were uncommon, I would have had no reason to write this article or the book from which it is drawn: *When Can You Trust the Experts? How to Tell Good Science from Bad in Education*. The field of education is awash in conflicting goals, research "wars," and profiteers. The goal of my new book is to help you evaluate new ideas related to education so that you are less likely to be persuaded by bad evidence, in particular, evidence that proponents claim is scientific.

The field of education is awash in conflicting goals, research "wars," and profiteers. My goal is to help you evaluate evidence that proponents claim is scientific.

Unfortunately, distinguishing between good and bad science is not easy. Evaluating whether or not a claim really is supported by good research is like buying a car. There's an optimal solution to the problem, which is to read and digest all of the relevant research, but most of us don't have time to execute the optimal solution. What we need is a good shortcut.

The shortcut I've developed is composed of four steps: strip it and flip it, trace it, analyze it, and make your decision about whether to adopt it.

"Strip it" means to lay the claim bare, devoid of the emotional language and other ornamentation that people use to cloak the actual scientific claim. Examining the claim in its simplest form can make many problems plain to you: the claim is true but self-evident, or the promised outcome is vague, or no one specifies the connection between what you're supposed to do and what is supposed to improve. "Flip it" addresses the fact that how we perceive the promised outcome is sensitive to the description provided; for example, saying that ham is "90 percent fat free!" sounds quite different than saying it is "10 percent fat!"

"Trace it" is applied not to the educational claim or program but to its inventor. Most of us use this step already and, in fact, overuse it. It means to pay attention to the qualifications and motivations of the person trying to persuade us. We are most convinced by people who are knowledgeable and impartial. Unfortunately, it's hard to judge whether or not someone is knowledgeable about a subject unless we ourselves have some expertise. We tend, therefore, to rely on credentials. We believe doctors when they speak about medicine, and electricians when they talk about our fuse box. Credentials can be faked, but even when they are genuine, credentials are not a reliable guide to believability in education. In fact, this most commonly used earmark of credibility is the *least* useful.

"Analyze it," the third step of the shortcut, means to consider why you are being asked to believe something. If the claims about an education product fly in the face of what you know to be true, there is a problem. At the same time, your experience is not an infallible guide. If it were, there would be no need for scientific research. So, "analyze it" also means to apply some simple guidelines to evaluate research claims. The point of the shortcut is to save you from having to evaluate research, so I don't suggest getting too technical here. But there are some useful rules of thumb to apply (like making sure a study that purports to show a program's effectiveness has both a treatment group that used the program and a comparison group that used something else).

After evaluating an idea's scientific merit, you need to decide whether or not it should be adopted. Although I'm advocating for a shortcut, I'm not advocating that a decision be rash. Nor am I saying that one should never adopt an educational program that lacks scientific support: most lack such support. What I'm arguing for is adopting a program only when you have all of the relevant information before you.

The shortcut is designed to help you evaluate the likely scientific soundness of a proposed curriculum, teaching strategy, textbook—anything that is purported to help children learn. Note that I said the *likely* scientific soundness. I freely admit—no, I emphasize—that what I'm recommending is not a substitute for a thoughtful evaluation by a knowledgeable scientist. Rather, it's a workaround, a cheat. As such, it's imperfect. The great advantage

is that it doesn't require a knowledgeable scientist.

In this article, I'll provide some detail on the first of the four steps: strip it and flip it. As a shorthand, I'm going to use the term *change* to refer to a new curriculum or teaching strategy or software package or school restructuring plan—generically, anything that someone is urging you to try as a way to better educate kids. I will use the term *persuader* to refer to any person who is urging you to try the change, whether he or she is a teacher, administrator, salesperson, or the president of the United States. To get started, you need to be very clear on three points: (1) precisely what change is being suggested, (2) precisely what outcome is promised as a consequence of that change, and (3) the probability that the promised outcome will actually happen if you undertake the change. All other considerations are secondary at this point and should be considered distractions.

This self-evident solution—take what works one place and implement it elsewhere—is a notorious flop. Successes depend on many factors that are hard to replicate.

Strip It

To strip a claim to its essentials, I suggest that you construct a sentence with the form "If I do X, then there is a Y percent chance that Z will happen." For example, "If my child uses this reading software an hour each day for five weeks, there is a 50 percent chance that she will double her reading speed." Of course, the agents might vary: the person doing X might be a student, a parent, a teacher, or an administrator, and the person affected by the outcome (Z) might be any of those. Note that the value of Y (the chance that the desired outcome will actually happen) is often not specified. That's fine. Right now all you're trying to do is be clear about the claim made by the persuader, and if she has left Y out, she's left Y out.

The purpose of stripping a claim is to remove cues that might be persuasive, even if they don't provide any real information. One such cue is an emotional appeal.

Stripping Emotion

The "If X, then Y percent chance of Z" formula will eliminate emotional appeals, which can be very powerful, indeed.

Emotional stories may add personal texture to a problem that we understood only abstractly, or make a problem seem more urgent, but they don't provide compelling reasons to do any particular thing. Why? Because emotional appeals don't provide evidence that a particular solution will work.

Persuaders in education seek to rouse different emotions, depending on their audience. For administrators and policymakers, it's most often fear. For example, consider these quotations from a column written by *New York Times* columnist Thomas L. Friedman in 2009:[9]

> Just a quick review: In the 1950s and 1960s, the U.S. dominated the world in K–12 education. We also dominated economically. In the 1970s and 1980s, we still had a lead, albeit smaller, in educating our population through secondary school, and America continued to lead the world economically, albeit with other big economies, like China, closing in.
>
> There are millions of kids who are in modern suburban schools "who don't realize how far behind they are," said Matt Miller, one of the authors [of a recent study]. "They are being prepared for $12-an-hour jobs—not $40 to $50 an hour."
>
> We urgently need to invest the money and energy to take those schools and best practices that are working from islands of excellence to a new national norm.

The persuader refers to broad economic trends and extrapolates a dark picture to the near future. Foreign, better-educated kids are in America's rearview mirror, gaining fast, and economic ruin will follow when they pass us. Fear makes us more open to suggestion: "That sounds terrible! Quick—tell me how to fix it!" But in fact, the message mentions a solution only briefly—invest money to take best practices from one school and put them in another—and provides no supporting evidence that this measure will work. In fact, this self-evident solution—take what works one place and implement it elsewhere—is a notorious flop among those who know the history of education policy. Successes depend on many factors that are hard to identify, let alone replicate.

When persuaders target teachers, they more often use emotional appeals centering on hope, not fear. Most teachers are optimists. They believe that all children can learn and that all children have something to offer the classroom. Teachers are also optimistic about the possibility that they can help children fulfill their potential. But teachers are not optimists to the point that they are out of touch with reality. A teacher knows when there is a child with whom she is not connecting. She knows if some aspect of her teaching has become grooved, familiar, and a little stale. When they talk to teachers, persuaders offer a change as a way finally to reach that unreachable child or to put the passion back into the teaching.

Administrators often try to sell teachers on an idea by dangling hope before them. Administrators know that "buy-in" is vital—if teachers don't believe a change is a good idea, they won't implement it in their classroom. Thus, administrators see a need not merely to persuade teachers, but to inculcate zeal for the change.

Fear does not encourage zeal. It encourages grudging compliance. Hope breeds zeal. That is why professional development sessions sometimes feel like evangelical revival meetings. But hope, like fear, is not a reason to believe that a change will work.

Stripping Claims that the Persuader Is "Like You"

When you change a persuader's claim to "If I do X, then there is a Y percent chance that Z will happen," the emotional language ought to vanish. So too should another set of irrelevant cues that might nudge you to believe something: those primed to make you think the persuader is *like you*, because we are, indeed, more likely to believe people we think are similar to us. Many websites and professional development marketers will claim quite directly, "I know what it's like…" The developer of the product will go to some pains to make clear that she's a teacher or a mom. Consider this example, from a website touting a treatment for attention deficit hyperactivity disorder (ADHD): "Your friends think he just needs consistency. Your doctor wants to medicate him. Your husband doesn't see why you can't control him. Your mom thinks he just needs a good spanking." By predicting the reactions of friends and family—reactions that would make a mom feel guilty or inadequate—the author signals, "I know what it's like to be you."

But being "like me" doesn't really increase the chances that you've got a solution to the problem I face. Lots of people "know what it's like" and haven't found an easy path to reading comprehension or a way to motivate frustrated kids or a method to help children with autism connect with other kids. And let's face it: being similar to your audience is an easy credential to inflate. I once attended a professional development seminar in which the speaker told story after story of his experiences in the classroom, all of which were, in turn, funny or poignant, and all of which showed that he "got" teachers. I later learned that he had been a classroom teacher for one year, 20 years earlier. He'd been doing professional development ever since, telling, I suppose, the same set of classroom stories.

Stripping Analogies

Stripping claims also removes the potentially powerful and often misleading role of analogies. When analogies are suggested to us, we tend to use them. That's why politicians so frequently offer analogies to defend their policies. For example, analogies were rampant in the United States during the buildup to the Persian Gulf War. Those who favored intervention drew an analogy between Saddam Hussein and Adolf Hitler: both were dictators of militaristic countries with regional aspirations who invaded weaker neighbors. Most Americans think that earlier action against Hitler could have saved many lives, so if Saddam is like Hitler, military action seems to make sense. But other politicians countered with a different analogy. Iraq is like Vietnam. Both were distant lands that did not directly threaten the United States. Most Americans regret the Vietnam War, so this analogy suggests *not* undertaking military action.

You would think that people would not be taken in. Surely we make judgments based on the merits of the case, not based on a rather shallow analogy suggested by a politician. But experimental data show otherwise. In one study, subjects read a fictional description of a foreign conflict and were asked how the United States should respond, using a scale from 1 (stay out of it) to 7 (intervene militarily).[10] The description they read did not explicitly offer an analogy, but instead dropped hints that were to make subjects associate the scenario with either World War II or Vietnam: for example, the president was said to be "from New York, the same state as Franklin Delano Roosevelt," or "from Texas, the same state as Lyndon Johnson." Later, they were asked to judge how similar the fictional scenario was to each of these conflicts.

There were two fascinating results in this study. First, people *were* influenced by the hints. People who read the story with the World War II hints favored intervention more than people who read the same story with the Vietnam hints. Second, people *thought* that they weren't taken in by the analogy. Both groups said that the story they read was not very similar to World War II and not very similar to Vietnam. In short, people thought, "I see how you're trying to influence me, but I'm too smart for you. The analogy you're suggesting doesn't really apply." But their judgments of how to respond showed that they were influenced nevertheless.

Analogies are sometimes offered in discussions of education, and that's another reason to strip claims. Consider this snippet adapted from a talk to a school board, similar to many that I've heard in the last five years.* The speaker was there to talk about the role of new technologies in education. Students today carry phones with more computing power than the desktop machines of 10 years ago. Many students are in contact with friends via social networking sites and text messages literally during every waking hour. What do those facts imply for education? Here's the nub of the speaker's argument:

> Let's consider what these new technologies have meant for various industries. Magazine publishing is almost defunct, and newspapers are desperately playing catch-up, trying to figure out a way to adapt. Remember those drive-up places to get your film developed? Remember stores that rented movies? Those are gone. People no longer use travel agents. They no longer use maps.
>
> All of these industries are obsolete, unnecessary. And they all have something in common; each was based on the delivery of information. These industries no longer exist because the Internet offers personalized, immediate access to almost limitless information.
>
> So what does that mean for schools? *Education is in the business of delivering information*. The pattern in other businesses has been for information delivery to become more mobile, real-time, and collaborative, and also to be more personalized. The question for teachers and administrators is, "How are you going to adapt?"

The speaker's message was clearly emotional—he was quite literally suggesting that everyone in the audience was going to be as obsolete as a VHS video player, and soon. But this suggestion was by analogy. Obviously, he's right when he says that various industries have been rendered irrelevant by new technology. But it's not obvious that every industry that delivers information is doomed. Education differs from these other industries in that a personal relationship (between teacher and student) is known to

*This example, like many I use, was inspired by a real talk, but I've changed it enough that it's not clearly attributable to the original speaker.

be central.[11] I don't need a personal relationship with the person who makes my airline reservation.

Other peripheral cues will also disappear when you strip a claim. Persuaders naturally want to appear authoritative. They will brag about academic degrees (if they have them). They will claim associations, however tenuous, with universities, especially prestigious ones, or they will claim to have consulted with Fortune 500 companies. They will boast about the authorship of books and articles; they will boast about speaking engagements. These are all indirect ways of saying, "Other people think I'm smart." They are not claims about the efficacy of the change, but rather are claims about the persuader. I go into greater detail about how to evaluate the persuader in my book, but here's a preview: characteristics of the persuader are a very weak indicator of scientific credibility. Stripping the claim will help you ignore them.

Flip It

Psychologists have long been interested in how people make decisions. We might bet that decision making is a complex cognitive process, but we'd also bet that certain things about that process can be taken for granted—for example, that the particular way you describe the decision I have to make shouldn't influence what I decide to do, provided that both descriptions are clear. That perfectly reasonable assumption turns out to be incorrect. People *are* affected by the description of the choice they are to make.

Flip Outcomes

Consider this: in one study, subjects were asked to sample cooked ground beef and were told either that it was "75 percent fat free" or that it was "25 percent fat." Subjects in the former group rated the beef as better tasting and less greasy.[12] This is one example of a large family of phenomena psychologists call *framing effects*. In framing effects, the way a problem or question is described influences the solution or answer we provide. This is why when you hear about an outcome (that's Z in our strip it formula), it's worth thinking about flipping it.

How might this be relevant to education? Just as a grocer would prefer to tell you how lean beef is rather than how fat it is, a persuader would rather tell you how many children will be reading on grade level if you adopt her change, and would rather not talk about the converse—how many will not. Although such framing seems like an obvious ruse, experiments show that providing information about success rates rather than failure rates actually makes people rate programs as more successful.[13] So when you hear that a curriculum promises "85 percent of children will be reading on grade level," flip it. Recognize that 15 percent won't. This failure rate may seem acceptable, but it's worth having it clear in your mind (especially since, if you implement this program, you'll need to find something else that is likely to be effective with the remaining 15 percent of children).

Flip What You're to Do

Another somewhat obvious framing effect doesn't concern the outcome (Z in our strip it formula) but rather concerns what you're asked to do (X in the strip it formula). Sometimes a problem is presented as though it is inevitable that we must take action. After all, there's a problem! Something must be done! But inaction is not always the worst possible choice. Years ago, a dentist told my father that his teeth were in terrible shape. He took about five minutes frightening my dad with all the details, and then another five describing an elaborate set of measures he might take to delay the inevitable, ending with, "Now if I do all that, I think you can keep your teeth for another ten years." So Dad asked, "Okay, what if I don't do any of that stuff. How long would my teeth last?" The dentist was taken aback that anyone would consider such a plan, but Dad persevered, and finally squeezed an answer out of him: "I don't know. Ten years, maybe?"

There are many problems in education with a similar profile: they are real problems, but there is no proven method of dealing with them. Thumping the table and insisting "Something must be done!" misses the point. Yes, lots of kids don't know as much civics as they ought to.[14] That doesn't mean we should plunge ahead with any civics program that we happen to lay hands on. Do we have some reason to believe that the new program will not make things worse? Is there reason to think that things might get better if we were to take no action? Or perhaps the "cure" being offered will avoid some problems but make others still worse. For example, some critics argue that children with ADHD should not be given medication. I understand the drawbacks: medications can have side effects, and the child may feel labeled by the diagnosis. Stopping the medication may solve those problems, but it incurs other costs; kids with untreated ADHD are at greater risk for dropping out of school, teen pregnancy, drug abuse, clinical depression, and personality disorders.[15] So here's another way to flip the persuader's claim: ask yourself, "What happens if I *don't* do X?"

Flip Both

A final framing effect is somewhat less obvious; to counteract it, you need to combine the two flips we've discussed. This won't seem as complex once we make it concrete, so let's start with an adapted version of the problem used in the classic experiment on this phenomenon.[16] Imagine that an island nation of 600 people is preparing for the outbreak of a deadly disease. There are two alternative medicines that can be used to fight the disease, but the constraints of time and money mean that the islanders can select only one. The scientific estimates of the medicines are as follows:

> **Medicine A**: 200 people will be saved.
>
> **Medicine B**: there is a one-third probability that 600 people will be saved, and a two-thirds probability that no people will be saved.

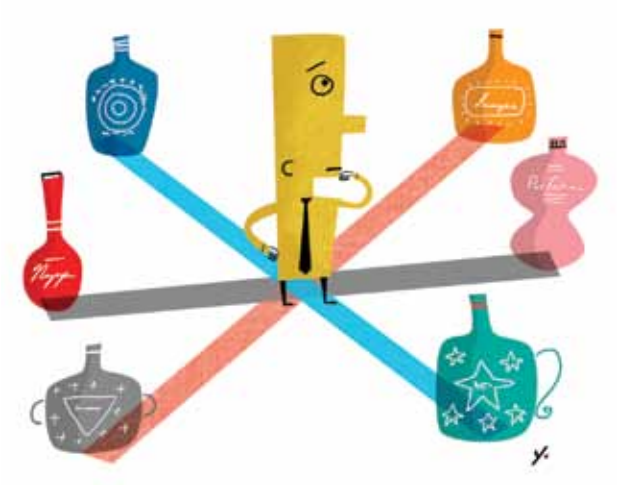
Which of the two programs would you favor? Before you answer, you should know that in this experiment, some subjects saw the version above, while others saw the same problem, but with a different description of the medicines:

> **Medicine A**: 400 people will die.
>
> **Medicine B**: there is a one-third probability that no people will die, and a two-thirds probability that 600 people will die.

Notice that medicines A and B have the same consequences in the two versions of the problem; "200 people will be saved" is the same outcome as "400 people will die." So now, like the hamburger situation (lean versus fat), we vary the description of the outcome (people saved versus deaths); but unlike the hamburger

situation, there's a choice to be made (rather than just rating the appeal of the burger).

The findings were striking. When offered the first description—which emphasizes the people saved—72 percent chose medicine A. But when offered the second description, which emphasizes deaths, just 28 percent chose medicine A. Why? Most psychologists interpret this as part of a very general bias in how we think about risk and outcomes. We are *risk averse* for gains, and *risk seeking* for losses. That means that when we must make a choice between two good outcomes (where we stand to gain something), we like a sure thing. Hence, when the medicines are described in terms of lives saved, we go for the sure thing—100 percent chance that 200 people will be saved. But, when losses are salient, suddenly we're ready to take risks to reduce the loss. Hence, in the second problem description, people are apt to choose medicine B, hoping for the outcome where no one dies.



Now let's put this into the strip it formula. In the first flip, I asked you to think about whether there is another way to describe the outcome (Z)—that's the lean versus fat hamburger business. In the second flip, I asked you to compare the outcome of adopting the change (X) to the outcome when you do nothing (not X), as in my dad's dentistry experience. In the island disease problem, we've combined them. Everyone was asked to consider a choice of what to do (X), but the outcome was described positively or negatively (Z).

Let's put this into an education context. Suppose you're a school principal and the central office in your district closely monitors the percentage of kids who read at grade level, as defined by a state-mandated test. With your current reading program, 34 percent of kids in your school are reading at or above grade level and 66 percent are not. If you adopt a new reading program, there is some chance that it will work well and things will improve. But there is also some chance that things will get worse—teachers will be unfamiliar with the new program and so won't implement it effectively, or the program just may not be as good as what you're doing now. We can frame this choice in terms of losses:

> **Choice A** (keep doing what you've been doing): 66 percent of kids read below grade level.
>
> **Choice B** (adopt new program): there's a two-thirds chance that 90 percent of kids read below grade level, and a one-third chance that 10 percent of kids will read below grade level.

Or we can frame the choice in terms of gains:

> **Choice A** (keep doing what you've been doing): 34 percent of kids read at or above grade level.
>
> **Choice B** (adopt new program): there's a two-thirds chance that 10 percent of kids read at or above grade level, and a one-third chance that 90 percent of kids read at or above grade level.

Naturally, I've fabricated the figures in these choices, but I'm sure you get the point. When we think about adopting a change, we understand that there's some chance that it will help, but there is also some chance that it will not work or even make things worse. We can frame these possible outcomes either as gains or losses. *When things are described as losses, we are more likely to take a risk*. So when a persuader emphasizes again and again that things are really bad, what is she really saying? She's saying that the current situation means a certain loss! The persuader is egging you on to take a risk. When the island problem was described in terms of losses (deaths), people were more ready to go for a risky solution to try to minimize the losses. If the persuader instead emphasized *gains*, you would be more likely to stick with what you're doing—where your gains are certain—rather than taking a risk to try to increase your gains.

> The particular way you describe the decision I have to make shouldn't influence what I decide to do, but people *are* affected by the description of the choice they are to make.

Whether or not the risk is worth it is, of course, a matter of the odds of the gains and losses, as well as how good the gains seem to you and how bad the losses seem. I'm emphasizing that you should look at these outcomes from all possible angles, because your willingness to try something risky is influenced by whether you think of yourself as trying to get something good or trying to avoid something bad.

Stripped, Flipped, and Clearly Not Worth Your Time

This first step in the shortcut—strip it and flip it—is meant to be devoid of evaluation. You are simply to gain clarity on the claim. One benefit of gaining clarity is that you can see that some claims are unworthy of attention. Once stripped and flipped, some claims are familiar, some are unacceptably vague, and some are so extravagant as to be unlikely to affect students. Let's look at each of these.

Familiar Stuff

One possibility is that the claim, once stripped of fluff, is revealed as something humdrum because it is already familiar. This phenomenon is especially prevalent in so-called brain-based education. Neuroscientific terms seem so impressive, so unimpeachably *scientific*, that it may not occur to you that the findings, though perfectly true, don't really change anything. The table below shows some neuroscientific findings that I have seen emphasized in books and blogs.

Neuroscientific Finding	Stripped
Dopamine, a neurotransmitter associated with both learning and pleasure, is also released during video gaming. Video games may be an ideal vehicle through which to deliver educational content.	Kids like games, so if we could make learning more like games, kids would like learning.
Although the brain weighs just three pounds, it commandeers about 20 percent of the body's glucose—the sugar in the bloodstream that provides energy. When glucose in the brain is depleted, neural firing is compromised, especially in the hippocampus, a structure vital to the formation of new memories.	A hungry child won't learn very well.
The prefrontal cortex of the brain is associated with the highest levels of decision making and rational thought. It is also the last part of the brain to be myelinated—that is, to be coated in the insulation essential to effective neural functioning. The prefrontal cortex may not be fully myelinated until 20 years of age.	Sometimes teenagers do impulsive things.
There is massive brain plasticity during the early years of life. Brain plasticity is the process by which the physical structure of the brain changes, based on experience. New networks are formed, and unused networks are "pruned" away—that is, are lost.	Little kids learn a lot.

Vague Stuff

Some claims, while far from mundane, are very hard to size up because they don't yield to your best efforts to put the claims into the format "If I do X, then there is a Y percent chance that Z will happen." In other words, you can't quite figure out either what you're supposed to do (X) or what is supposed to happen after you do it (Z). That problem ought to strike you as quite serious. You are embarking on this educational change because you think it's going to do some good. If you don't have it clear in your mind what Z is supposed to be, then you can't know whether or not the change is working. And if you don't have X clear in your mind, that means you're not sure whether you're doing the right thing to make Z happen.

Take, for example, the change of placing an interactive whiteboard* in a classroom. It would seem that this tool could be quite useful in a classroom. For starters, the teacher can capitalize on all of the software on the web. The United Kingdom invested heavily in interactive whiteboards, and today virtually every UK school has at least one. But the impact on student achievement has been minimal. It turns out that the presence of an interactive whiteboard in the classroom does not necessarily change teaching for the better, or even change teaching at all.[17] Teachers need not only the whiteboard but also substantive training in its use, expert advice about how to exploit it in lesson plans, and time to gain expertise and confidence (all of which, if it were provided, would fill in X in our formula).

It's not just technological changes that are underspecified. Many changes that urge project learning or group learning have this characteristic. Just as dropping an interactive whiteboard into a classroom is not enough to ensure that students will learn, assigning group work is not enough to ensure that students will learn how to work well in groups. These pedagogical approaches call for much more independence on the part of students, and therefore they depend on the teacher having strong relationships with the students and a good understanding of the existing relationships between students. The teacher uses this knowledge in hundreds of moment-to-moment decisions that guide the groups in the work without micromanaging them. Thus, changes that suggest lots of group work in the classroom are almost always underspecified. The methods are terrific when they work well—in fact, I think that for some types of learning they are probably ideal—but they are very difficult to implement well, and I seldom see a persuader acknowledge this difficulty.

The clarity of the outcome is just as important as the clarity of what you are supposed to change. For example, suppose that my son's first-grade teacher has told me that he's struggling with reading, and I notice that he shows no interest in reading at home. I hear about a technique called Language Experience[18] that is supposed to help struggling readers, and I decide to give it a try. Language Experience is quite specific about what you're supposed to do:

1. You have the student dictate something to you (a story, a description, anything that the student would like to relate).
2. You write down what the student says, periodically stopping and reading aloud to the child what you have written so far.
3. When the child is finished, you read the whole piece aloud to him or her.
4. You save the piece so that the child can reread it himself or herself.

The method is clear enough. The outcome, less so. It's supposed to help make reluctant readers more interested in reading. Okay, but how are you to know that's happening?

Knowing what a change is supposed to do is not quite the same as being able to evaluate whether or not it's actually happening. If a persuader promises that a change will make kids like reading more, how will I know that they do? I could just ask them: "Do you like reading more than you did six weeks ago?" But then again, maybe children's memory for that sort of thing is not that accurate. Then too, if the child says, "Yes, I like reading more," but then seems just as miserable during reading time at school, should I be persuaded by what she says, or by how she seems to act? If I am to evaluate whether a change is working, I need something concrete, and something that is well matched to what I was hoping the change would do. For example, perhaps I was prompted to look for a reading program because my child complained about reading in school and seldom read books at home; I could see

*An interactive whiteboard is used as a screen on which one can project an image from a computer. The screen is touch sensitive, so the teacher (or student) can interact with the computer by touching the screen.

whether the change prompts less complaining and more reading.

I also need some idea of what constitutes "success." Suppose that in the week before my son starts this new reading program, he doesn't pick up a book once. If, three weeks into the program, he is looking at books once each week, am I satisfied? Or does that change seem too small? In addition, I need to know when to expect that the good outcome will have happened. For example, you'd think it pretty odd if I told you that I had been using a reading program for two years with no sign of it helping, but I was still hopeful that eventually it would do some good. Okay, so two years without results is too long. What's more reasonable? Two weeks? Two months?

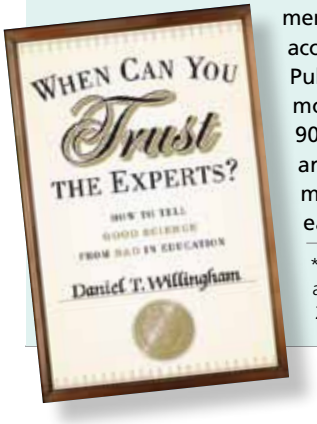
Once you've invested your time, you're likely to persuade yourself that the program is working, even if you're grasping at straws.

It's important to define the signs of success *before* you embark on the change. Once you're committed, your judgment of how it's working is all too likely to be affected by cognitive dissonance. "Cognitive dissonance" refers to discomfort that is a consequence of holding two conflicting beliefs simultaneously—and it may make it hard to evaluate how an educational change is working. Once you have been embarked on a change for a while, you've invested your time and that of the students or your child, and you may have a financial investment. Thus, if the change isn't really working that well, you will hold two incompatible thoughts in mind: (1) I invested heavily in this program, and (2) this program brings no benefits. It's hard to rewrite history and pretend that you *haven't* invested in the program, so you are likely to seek out reasons to persuade yourself that the program is working, even if you're grasping at straws.

The best way to protect yourself from this profitless self-delusion is to write down your expectations before you start the program: how big a change you're expecting, when you expect to see it, and how you'll know the change is happening. Writing down these expectations makes it difficult for you to persuade yourself that something is working when it's not, because you have already defined for yourself what it means for the change to be "working."

When Can You Trust the Experts?

Suppose you're a doctor. You go through medical school and residency, learning the most up-to-date techniques and treatments. Then you go into family practice, and you're an awesome doctor. But science doesn't stand still once you've finished your training. You were up to date the year you graduated, but researchers keep discovering new things. How can you possibly keep up with the latest developments when, according to PubMed.gov, more than 900,000 articles are published in medical journals each year?*



*PubMed.gov, accessed June 10, 2011.

Medicine has solved this problem for practitioners by publishing annual summaries of research that boil down the findings to recommendations for changes in practice. Physicians can buy summary volumes that let them know whether there is substantial scientific evidence indicating that they ought to change their treatment of a particular condition. In other words, the profession does not expect that practitioners will keep up with the research literature themselves. That job goes to a small set of people who can devote the time needed to it.

In education, there are no federal or state laws protecting consumers from bad educational practices. And education researchers have never united as a field to agree on methods or curricula or practices that have sound scientific backing. That makes it very difficult for the nonexpert simply to look to a panel of experts for the state of the art in education research. There are no universally acknowledged experts. Every parent, administrator, and teacher is on his or her own. That's why I wrote this book.

This book will not turn you into a research expert. Indeed, the point of the book is to obviate the need for expertise. And the shortcut I offer is imperfect, like all heuristics. You might apply these methods and still draw the wrong conclusion. But I can promise this. Whatever your current level of research sophistication, this book will help you ask better questions about the research base behind a product, and it will help you think through the wisdom of purchasing and using a product in your classroom, school district, or home.

–D.T.W.

Extravagant Stuff

Some claims about changes are neither familiar nor vague; they are too extravagant. From a cognitive perspective, if a persuader makes either of the following two promises, they are very unlikely to be kept: (1) that a change will help with *all* school subjects, or (2) that a change will help *all* kids with a particular problem. Let's consider each in turn.

Suppose that instead of being tutored in academic subjects, students performed a set of exercises tapping basic mental processes that underlie *all* cognition. You don't just tutor the student in history; instead, you make memory work better, or you improve critical thinking. Many of the "brain games" software packages and cognitive training centers make such claims.

The problem is not just that you can't train basic cognitive processes like working memory. The problem is that when you practice a cognitive skill—critical thinking, say, or problem solving—the newly acquired skill tends to cling to the domain in which you practiced it. That is, learning how to think critically about science doesn't give you much of an edge in thinking critically about mathematics. There are two reasons that critical thinking sticks to subject matter: sometimes you need subject knowledge to recognize what the problem is in the first place, and sometimes you need subject knowledge to know how to use a critical-thinking skill.[19] So when I see a change promise to improve a skill (such as "critical thinking") and it makes no mention of the need for knowledge to go with it, I'm suspicious.

The second type of across-the-board claim that ought to make you leery does not cut across the cognitive abilities of one child, but rather concerns a single ability in many children. I am suspicious of changes that promise to remediate a problem in *any* child. Why? Because each of the outcomes we care about for schooling is complex. Lots of cognitive and noncognitive processes contribute. Put another way, if a child is having problems with reading, there are many possible reasons for that. Thus, a change might help with reading difficulties that are due to a problem in processing sounds, but that's not going to work for a child who has a problem with visual processing. Hence, when a persuader claims that a change will help *any* child with a reading difficulty, the needle on my nonsense detector flutters close to the red zone.

We've covered the first of four steps in my shortcut for evaluating claims about educational changes. The table below summarizes all of the subcomponents of step one: strip it and flip it.

I urge you not simply to think about the actions in the table below but to write down your thoughts about them when you are considering a change. Forcing yourself to write things down will

(Continued on page 40)

Summary of Strip It and Flip It

Suggested Action	Why You're Doing This
Strip to the form "If I do X, then there is a Y percent chance that Z will happen."	To get rid of emotional appeals, peripheral cues, and proffered analogies that may influence your belief. The scientific method is supposed to be evidenced based and uninfluenced by these factors.
Consider whether the outcome (Z) has an inverse; if so, restate the stripped version of the claim using the inverse.	To be sure that you appreciate all the consequences of the action—for example, that an "85 percent pass rate" implies a "15 percent failure rate." We are subject to framing effects; we think something is better if the positive aspects are emphasized rather than the negative.
Consider the outcome if you fail to take action X.	To ensure that the promised outcome if you do X seems much better than if you don't do X. When there is a problem, it's tempting to lunge toward any action because it makes you feel that you are taking some action rather than standing idle.
Consider the outcome if you fail to take action, this time using the inverse of Z as the outcome.	To ensure that doing something versus doing nothing looks just as appealing when you think about good outcomes as when you think about bad outcomes. People are generally less willing to take risks to increase their gains—they would rather have a sure thing (even if the certain gain is small). But they are willing to take risks to minimize losses.
Evaluate whether the stripped promise is something you already know.	To be sure that what's being sold to you is something you can't do yourself. Technical talk—especially neuroscientific talk—can make old ideas seem cutting edge.
Evaluate whether the change (X) is clear; "clear" means that you feel confident that you know what to do and how the change will affect students' minds.	To ensure that the change is implemented as intended. Changes that sound good can go awry if they are not implemented in the classroom as intended or if students don't do what you're hoping they will do.
Evaluate whether the outcome (Z) is clear; "clear" means that there is some reasonably objective measure of whatever outcome you expect, how big the increase (or decrease) in the outcome will be, and when it will happen.	To be sure you will be able to tell whether or not the promised outcome is happening.
Check the outcome against this list of frequently claimed but extravagant and unlikely-to-work promises.	To be sure that claims are not unfeasible from a cognitive perspective—for example: an improvement in all cognitive processes, an improvement in a specific cognitive process (for example, critical thinking) irrespective of material, or an improvement for all students who struggle with a complex skill such as reading.

Telling Good Science from Bad

(Continued from page 12)

make you take more time with each action, and articulating your thoughts will increase their precision. It's well worth the time now, given that a change usually represents a significant investment of your time, money, and energy, not to mention the time and energy of your kids. If you do take the time, you'll see that many changes do not stand up to being stripped and flipped. As we've discussed, some will be familiar, vague, or too extravagant. Others will lose all appeal once stripped—there was nothing persuasive about them without the emotional appeal or misleading analogy. And still others will not seem impressive enough to be worth the investment once flipped.

I believe that the practice of education would be improved if better use were made of scientific advances, and if educators were better able to discern good science from bad. Will we continue to cheer on education reforms that sound right to us, convinced that the "evidence" supporting them must be strong only because we like the conclusion? Or will we cast a cold eye on our own beliefs, confident that, to paraphrase Francis Bacon, by beginning with doubt, we will end with certainty? If we can do so, our children will be the richer for it. □

Endnotes

1. David Lazarus, "If Nothing Else, Man with Past Is Persistent," *San Francisco Chronicle*, March 10, 2002.

2. B. J. Dohrmann, "Whole Brain Learning," www.superteaching.org/stmind.htm.

3. Gina Hannah, "Bernhard Dohrmann," *Huntsville Times*, April 28, 2002, A9.

4. This figure is according to the Super Teaching purchase order: http://superteaching.org/CEO_ST_purchase_order_v4.pdf (accessed June 13, 2012; this PDF is no longer available).

5. Budd McLaughlin, "'Learning at the Speed of Thought,'" *Huntsville Times*, October 7, 2008, 1A.

6. J. Ramhold, "University Dissolves 'Super Teaching' Partnership," *The Exponent*, April 14, 2010, http://exponent.uah.edu/?p=2538 (accessed July 17, 2011; this web page is no longer available).

7. This blog entry is no longer available from the Flashpoint blog website (www.flashpointblog.com).

8. A. Shavers, "Super Teaching: Learning at the Speed of Con," *The Exponent*, October 21, 2009, http://exponent.uah.edu/?p=1570 (accessed July 17, 2011; this web page is no longer available).

9. Thomas L. Friedman, "Swimming Without a Suit," *New York Times*, April 21, 2009.

10. Thomas Gilovich, "Seeing the Past in the Present: The Effect of Associations to Familiar Events on Judgments and Decisions," *Journal of Personality and Social Psychology* 40, no. 5 (1981): 797–808.

11. See, for example, Bridget K. Hamre and Robert C. Pianta, "Early Teacher-Child Relationship and the Trajectory of Children's School Outcomes through Eighth Grade," *Child Development* 72, no. 2 (2001): 625–638.

12. Irwin P. Levin and Gary J. Gaeth, "How Consumers Are Affected by the Framing of Attribute Information Before and After Consuming the Product," *Journal of Consumer Research* 15, no. 3 (1988): 374–378.

13. See, for example, Mark A. Davis and Philip Bobko, "Contextual Effects on Escalation Processes in Public Sector Decision Making," *Organizational Behavior and Human Decision Processes* 37, no. 1 (1986): 121–138; and Kenneth J. Dunegan, "Image Theory: Testing the Role of Image Compatibility in Progress Decisions," *Organizational Behavior and Human Decision Processes* 62, no. 1 (1995): 79–86.

14. National Center for Education Statistics, *The Nation's Report Card: Civics 2010* (Washington, DC: U.S. Department of Education, 2011).

15. Russell A. Barkley, *Attention-Deficit Hyperactivity Disorder*, 2nd ed. (New York: Guilford Press, 1998).

16. Amos Tversky and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science* 211, no. 4481 (Jan. 30, 1981): 453–458.

17. Julia Gillen, Judith Kleine Staarman, Karen Littleton, Neil Mercer, and Alison Twiner, "A 'Learning Revolution'? Investigating Pedagogic Practice Around Interactive Whiteboards in British Primary Classrooms," *Learning, Media and Technology* 32, no. 3 (2007): 243–256.

18. Elements of this technique go back quite far. One of the more influential presentations is Roach Van Allen and Claryce Allen, *Language Experiences in Early Childhood* (Chicago: Encyclopedia Britannica Educational Corporation, 1969).

19. For more on this, see Daniel T. Willingham, "Critical Thinking: Why Is It So Hard to Teach?," *American Educator* 31, no. 2 (Summer 2007): 8–19, www.aft.org/pdfs/americaneducator/summer2007/Crit_Thinking.pdf.

INTENTIONALLY LEFT BLANK