Chapter 2

Science and the Education of Teachers

JAMES M. KAUFFMAN
University of Virginia

ABSTRACT: *The preparation of teachers should be as scientific as we can make it. Preparing teachers in the scientific tradition requires embracing scientific and mathematical views of things that are impossible and those that are inevitable. It also requires knowing what science is and is not. Making teaching an applied science will require developing manuals and checklists for instruction and adopting the notion that teaching should become a manualized profession. All professions basing their work on scientific evidence and field tests develop manuals and checklists to guide their practices. Education must do the same.*

We must soon decide whether education—particularly teacher education—is going to be made significantly more scientific or continue very much as it has been practiced. In many ways, teacher education is at a crossroads. Teacher education's inadequacy was summed up well by Snider's (2006) description of her undergraduate training in education:

> I learned very little in my undergraduate teacher education program about how to teach; and for those first 8 years I relied on luck, trial and error, and the competence of colleagues for my professional development. I regret that I didn't know more from the beginning because despite my earnest efforts, my students didn't achieve as much as they could have. I knew very little about curriculum, effective teaching, or principles of classroom management beyond what I learned on the job. (p. 2)

This description is not very different from what many of us have experienced. Lack of training in direct instruction or Direct Instruction (see Kauffman, 2010, for a description of differences between di and DI) is understandable for those of us trained before these effective ways of teaching were described clearly. But why are most prospective teachers *not* taught to use DI today? DI is a readily available, scientifically and logically derived, field-tested program for teaching that, to my dismay, is still widely ignored. To me, this is inexcusable. We in higher education must do better. We must make education and the preparation of teachers an applied science (Kauffman, 2011) and a logically derived endeavor (Engelmann & Carnine, 2011). And we must do better than alternatives such as Teach For America and Troops to Teachers. If teacher preparation in higher education remains what it has been, it will very likely be completely discredited, as it probably should be.

This paper is divided into three major sections. The first section is about some impossibilities that people seem to hanker for but that just aren't possible—and then some inevitabilities that people tend to ignore. We ought to get over the fantasy that we can achieve the impossible or avoid the inevitable. I highlight impossibilities and inevitabilities because if we cannot get comfortable with them, then we have no real hope of making education a science. The second section describes a few of the many common misconceptions about science, with special attention on how they apply to education. Common misconceptions about science can lead us to false conclusions. The third section is about why we need manuals and checklists in education, especially in teacher education. It makes a case for teaching teachers to teach by the book and for using devices to help keep us from making common errors that are especially costly to learners.

IMPOSSIBILITIES AND INEVITABILITIES

Impossibilities

Some things just aren't possible. We know this because of some very basic realities of math and science. True, some things once considered impossible are possible today, and some of the things we consider impossible today might be possible some day. However, some things will always remain impossible, such as adding two positive integers and obtaining a sum less than either of them or talking about something without using a word or words for it. Unfortunately, some people either implicitly or explicitly assume that we do not really have to deal with unchangeable realities, that these realities can either be ignored or treated as inconvenient truths. For example, some people seem to think that universal proficiency, something impossible by definition, is achievable and raise questions only about the year in which we might reasonably expect it to be achieved. One newspaper editorial—without irony, obviously not noticing the mathematical impossibility, practical nonsense, or self-contradiction of its statement—praised the goal of 100% proficiency in 2014 set by the No Child Left Behind Act (NCLB), saying that the goal of universal proficiency "while laudatory, may be unrealistic" (*The Washington Post*, 2007).

To understand what I've said, you first have to think about what "proficiency" means. Proficiency in any skill (say, swimming or driving or math) is defined by what most people can do after specific training; it isn't a level of performance just pulled out of the air without reference to what people can do. Proficient/not proficient isn't a distinction based on what only a few of the highest performers can do, nor is it based on the performance of the most inept. So, to say that *all* people will do what *most* people can is simply a self-contradiction, a logical impossibility. It's as comical as saying that we're going to have all the children above average. Don't misunderstand. We can often help *more* people become proficient at something, but *all* people? Well, *universal* proficiency—*all* students becoming proficient in an academic skill, for example—just isn't in the cards. Getting more people proficient at something could be very hard but possible for some skills. Truly universal proficiency? No. Won't happen. Those of us who work with students who have severe disabilities understand that universal proficiency in reading, for example, just isn't possible. We don't approve of terms like "all" or "universal" being used as if our kids aren't considered. In education, it's important to think about the meaning of what we say and to say what we mean as precisely as possible. We want our students to do that. We should do that too.

Actually, the impossible is not a laudatory goal. That is, it's not good to set our sights on something that's logically, mathematically impossible. And it's not good to say things we don't mean. It's tempting to make sarcastic remarks about ill-considered comments like those in the *Post* or the contention that *all* children, regardless of their level of ability, should go to college or be prepared for a career—verbal equivalents of waving to Ray Charles (Kauffman, 2005). Real-world talk about education is more likely to better the lives of children than fantasy talk is.

Another impossibility is measuring something reasonably precisely without getting a statistical distribution. In education, this means a distribution of scores ranging from lowest to highest and having an average. Measuring educational performance accurately without getting a distribution with what statisticians call "moments"—mean, standard deviation, skew, and kurtosis, for example—is impossible. There are no exceptions. And this means that it is impossible to find that all of the individuals measured are at or above any location on that distribution except the lowest point. So, finding that all of the students are at or above the 20th percentile, for example, is impossible. Regardless of what a secretary of education or the United States Congress or someone with a Ph.D. hopes for, sets as a goal, or decides *should* happen, it is just not possible with the kind of mathematics we have on planet Earth. Consequently, NCLB was dead on arrival because it assumes that *all* students—or very nearly all, even excluding 2% or so of students who have disabilities—can be judged proficient by their state's test scores.

Now, probably I should explain a little about that 20th percentile statement I made. We can almost always improve students' performance. And, depending on the comparison we make, we could have more than 80% of students scoring above the 20th percentile—*of a distribution of test scores other than the one in which they were included*. That is, it's possible to have all of a particular group of students who took a test be above the 20th percentile of a different group of students who took the same test. For example, we could find that in a given school all of the students who took the SAT in a particular year scored above the 20th percentile on the SAT norm (i.e., above the 20th percentile of the group that took the test for norming purposes and established the 20th percentile for the norm). So, there's always the question of what comparison we want to make. Do we want to compare the students in the group that just took the test to each other on that test, or do we want to compare them to another group? Sure, we might get all students above a percentile greater than zero if the percentile refers to the percent in a different group, but not if it's the group we have. Let's think a little more about this.

It's possible to "play games" with statistics, even to play a game that makes something look good. Sometimes the game is played fairly. Comparison to an existing norm, perhaps even an old one, *can* make sense. But, suppose we want to make a group look good in such a comparison, even if it's sensible, by showing that everyone in the group is above the Xth percentile (i.e., any percentile greater than zero) for another group. We could do three things, and here's where the game gets really tricky and can be played to mislead people. First, we could choose a lower percentile; the lower the percentile, the greater our chance of getting everybody above it. Second, we could compare a smaller group to a larger group; the smaller the group we compare to a larger group, especially the normative group, the better our chances of getting everybody above a given percentile of the larger group. Third, we could make a comparison to a group that includes a lot of low performers; the greater the percentage of low performers in the comparison group, the better our chances of looking good by comparison.

Which reminds me of another thing we might consider: If we get all of the students above, let's say, the 20th percentile of some older test or normative group, then should we consider the older test outdated because the old norms aren't valid? The point is that we could make a comparison that isn't really sensible. And sometimes it's illogical, not sensible, to make a comparison to another group of test takers. But let's get back to the meaning of 20th percentile. For any given group that takes a test, we can't have more than 80% of the students who take that test above the 20th percentile of their group simply because the meaning of 20th percentile is that 20% of those who took the test got that score or a lower one. It's impossible to have more than 80% above the

20th percentile of that group for the same reason that we can't have more than 100% of a group.

Other examples of impossibilities that people call for can be found, like *an elite education for everyone*, which by definition is impossible (Kauffman, 2010). Garrison Keillor's description of Lake Wobegon is funny because we realize that it is impossible to have all of the children above average. We should not allow silly statements about children or schools, such as a goal of universal proficiency, to carry any legitimacy in serious talk about education. Lake Wobegon talk and the goal of universal proficiency are the stuff of comedy, not of serious thinking about educational outcomes.

Inevitabilities

Some things happen whether we want them to or not; they are inevitabilities, just facts of life that we should acknowledge and not think we can ignore. They are the flip side of impossibilities, realities that won't go away even if we wish they would. One example of the inevitable is the reality that some children are not going to learn to read. Ever. Even basic sight words. And some are not going to learn to read with what we consider fifth-grade comprehension. No matter what program we use or who teaches them.

What makes me especially angry is that many people do not include these children when *talking* about the education of *all* children—which, presumably, reflects their *thinking* about what is involved in teaching *all* children. These children, who don't "measure up" to the expectation that *all* children will reach a certain level of educational performance, are just written off, apparently. So when NCLB or some other misbegotten policy calls for *all* children to learn… whatever… these children are not even considered. They're assumed not to count, to be insignificant, and they and their teachers are assumed to be irrelevant or incompetent. As a special educator, I realize that there are children of school age who cannot walk or talk or communicate, cannot feed or toilet themselves, and need care and supervision around the clock. Yet these children are to be loved and respected and taught all the skills they can learn. But supposing that they can be made ready for college or a career is just preposterous. I also realize that there are children at every point on the distribution of ability.

The "bell curve" or normal distribution is often condemned, but the realities of statistical distributions of whatever we care enough about to measure will not go away (Kauffman & Lloyd, 2011). True, people sometimes make appalling assumptions about individuals related to distributions, but equally appalling is the assumption that the bell curve (i.e., a statistical distribution) can be ignored, called irrelevant in making policy decisions or training teachers, or simply wished out of existence. Although people may well be mismeasured (Gould, 1996b), regardless of the way they're measured we must consider what Gould

(1996a) called the "full house"—in education, all of the children to be taught including those far below average in whatever skill is measured.

Then there is the problem of prevention. I think the following bears special attention: Most people love the idea of prevention in the abstract, but they do not want to face its inevitabilities in practice. They seem to forget that it is impossible to prevent what has already happened. Prevention requires *anticipation* of whatever is supposed to be prevented (Kauffman, 2003). People often argue that we now misidentify many students as having disabilities but that we should practice prevention. Perhaps they do not understand that if many students are now misidentified as having disabilities, then prevention inevitably means that we are going to misidentify even more.

Imagine what would be inevitable if we actually practiced prevention. First, think about measuring whatever it is you want to prevent—maybe reading failure, maybe behavior problems, maybe something else. If you think about measuring it, then you are going to visualize the result—a distribution of scores that might approximate a normal curve. But even if the distribution you imagine is skewed (lopsided) or leptokurtotic (bunched up) or platykurtotic (spread out), think about what prevention requires. Prevention requires keeping as many individuals as possible from getting far from the central tendency of that distribution.

The basic idea of prevention is catching students earlier, before their problems get so bad. The idea is that if we catch a problem reader in first grade, for example, and we offer really effective instruction, then that student won't be so far behind come fourth grade. So, prevention requires moving the criterion or trip point for doing something about the problem (usually, we call this an intervention) closer to the central tendency. The distribution means, mathematically, that prevention requires including more individuals, not fewer, in the intervention. It requires increasing the risk of a false positive—identifying a child for help who doesn't really need it. Thus, complaints that we already serve too many students in special education and expressions of the unacceptability of misidentification are really arguments *against* prevention. Now, admittedly, if we move the criterion for receiving special education to more severe cases and leave prevention to the province of general education, then complaints that too many children are receiving special education can make some logical sense. However, then the argument for prevention becomes one of suggesting that more children should be identified earlier in general education, and also that the criterion for getting special education should require children to fail as much as they do now.

But for now, consider another problem that often upsets people when it comes to identifying children for special education or any other intervention—the mistakes in identification called false positives and false negatives. People generally don't like drawing a line, because it's arbitrary and some children are always close to it. The children who are in the close-to-cut-point (CTC) areas of a distribution are the ones who don't quite make it into the special program for some reason or who are in a special program when they perhaps shouldn't be. It is always possible to argue that some children were selected for a special program when they should not have been (false positives) and some should have been selected but were not (false negatives).

A line or criterion for a special program is an inevitable part of having a special program. Lines, labels, sorting—they are all necessary, inevitable aspects of having a special program. Who should get help with their mortgage? Having a program of financial assistance requires a line for qualification, a label describing those who participate in the program, and a means of sorting those who need help from those who do not. The same problems go with the designation of Tier 2 or Tier 3 in response to instruction or levels as does any other program in education that does not include every single child (Kauffman, 2010; Kauffman & Lloyd, 2011).

Another inevitability is that every line has margins (we might call them standard errors). No exceptions. Those who complain of misidentification may suggest multiple lines. But the problem is that every additional line creates more margins and more mistakes. Always. This is just basic math and science. Some cases are always CTC. Having three tiers instead of two in response to instruction increases the chance of making mistakes by about 50%.

Examples

We might consider an example of measuring reading ability and designating a level of performance that signifies reading failure. Remember, measuring without getting a distribution is impossible, and drawing a line or cut point for qualifying for a special program is inevitable. Figure 1 depicts a distribution of reading scores, with lower reading scores on the left of the curve. Possible cut points representing reading failure, A and B, are shown, along with their margins. If the cut point is moved from A to B (i.e., from more severe to less severe reading problems), then more students are included in the definition of reading failure. Moreover, by having two cut points, A *and* B, each a different tier, we double our chances of making an error—a false positive or a false negative. Two points about inevitability are noteworthy. First, moving the cut point toward less severe problems inevitably involves more individuals (i.e., a greater area under the curve). Second, every cut point has margins, its CTCs—areas of uncertainty on both sides of the cut point, simply because no test or other means of judgment is faultless, containing no error.

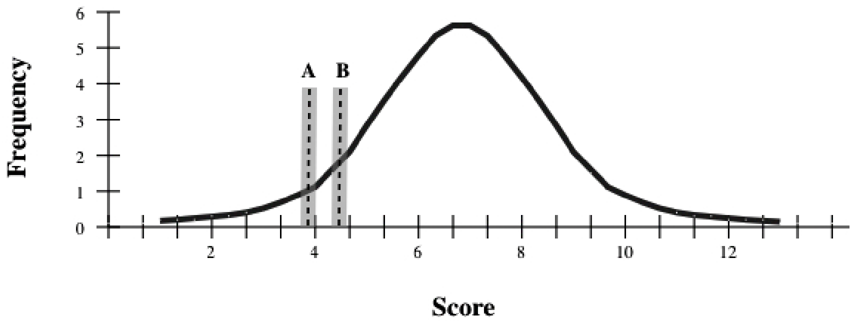

Figure 1. A hypothetical distribution of reading scores with alternative cut points A and B (and their associated margins or borders of uncertainty) indicating reading failure.

Figure 2 is a depiction of a curve of behavioral problems or disorders. It is another way of showing that as we move a cut point for defining disorder toward less severe cases (in this case, from A to B or C or from B to C), we inevitably include more children in our definition.

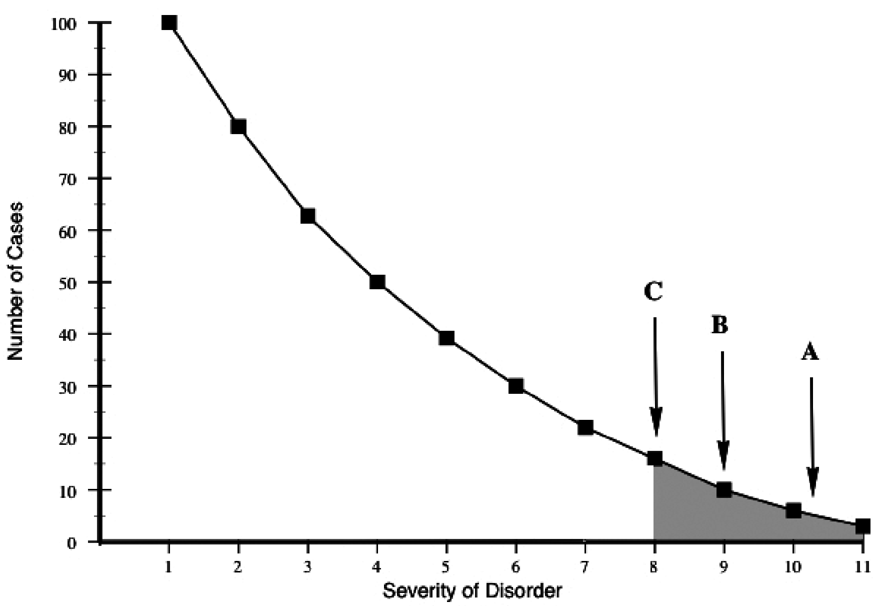

Figure 2. Number by serverity. Hypothetical curve of emotional or behavioral disorders showing alternative cut points A, B, and C defining disorder.

In no way does acknowledgment of the inevitabilities depicted by Figures 1 and 2 deny the agony of making difficult decisions that affect children's lives. However, we must acknowledge realities, like cut points and CTCs, in talking and writing about education. We have more than enough evasion of realities in rhetoric about education already. We do need precision in teaching, but we also need precision in our language and *thinking* about teaching. Perhaps we should start with careful thinking about what is possible and what is inevitable. The call for evidence-based practices must be consistent with reality-based thinking. "Never-never land" thinking does not help us or children.

The impossibilities and inevitabilities I have described are very fundamental ideas about realities, about how the world works, about mathematical functions that we cannot wish away. If we find it inconvenient to deal with them, then I think we have no real hope of adopting a scientific approach to education.

COMMON MISCONCEPTIONS ABOUT SCIENCE

It is very hard to get people who have not been trained as scientists to think scientifically about *most* things, including education (Kauffman, 2011; Landrum & Tankersley, 2004; Sagan, 1996; Specter, 2009). A very common belief is that teaching and learning cannot be researched in a scientific way. Some will argue that even if certain aspects of teaching can be made a science, the scientific research on teaching is trivial or meaningless—that the truly important things that happen between teachers and students, especially in students' minds, just can't be a matter of scientific study with important implications for teaching.

True, a science of education is very difficult, but it is not impossible. A science of education is in many respects more difficult than the "hard" or "bench" sciences because the kinds of *control* that can be achieved in most physics experiments and the kind of *stability* that characterizes earth science are impossible in education. Berliner (2002) was correct in saying that a science of education is the hardest science of all. Some scientists say that education is not a science at all. And to some extent, they're correct because education is not reliably scientific now and never has been. But I do not agree with the scientists who say that education is beyond the *reach* of science, for reasons that I hope will become clearer as you read this section of my paper.

A science of education is extremely unlikely without an understanding of what science is and how it works in the general case. Therefore, in this section I review a few of the basic principles that apply to *any* science and give some examples of how these principles might apply to education. (For a discussion of more principles than those covered here and a fuller treatment of science in education, see Kauffman, 2011.)

A serious science of education could disenthrall us from magical thinking about impossibilities and inevitabilities. It could help us find out *what* works and help us understand *why* something works or the *conditions* under which it works and doesn't work. Following the path of science does not guarantee that we won't make any missteps. In fact, if we take a scientific view of education, we'll make some mistakes. However—and this is important—our being scientific means that we will eventually, if not immediately, recognize our errors. As my friend, biologist Dan Burke, commented, "Science is not a steady parade of 'truth' but more a tortured path of six steps forward and five steps back, but generally moving in the right direction" (personal communication, December 27, 2009).

Science is commonly misunderstood. Not just by teachers, but by the general public. People would often rather be illiterate or disbelieving when it comes to science, even in areas like medicine (Sagan, 1996; Specter, 2009). We educators face a tremendous challenge in trying to help people understand how science might be applied to teaching and the advantages of a scientific approach to instructional problems.

Judgment Versus Certainty

One of the first things to come up in talk of science and education is *judgment versus certainty*. Contrary to popular opinion, data do not speak for themselves. Scientists must speak for data to make sense of their findings. Educators must use their judgment to urge action based on what they see as the *preponderance* of evidence rather than unarguable results. For example, DI ought to be supported because the preponderance of evidence suggests it is more effective than whole language.

Disproof Versus Proof

Scientists understand the idea of *disproof versus proof*. Science is not really the pursuit of direct proof but of things that can't be disproved—indirect proof. Scientists try to find something wrong with findings or explanations, and if they can't find anything wrong, then accept what they have found as their best guess—a tentative truth (Baldwin, 2008). In education, we very seldom can assume that something has been proved. More often, it's possible to draw the conclusion that something has been *disproved* beyond any reasonable doubt. A given procedure may be shown *not* to produce the desired result. So, then, we conclude it's no good or doesn't work. If something is shown *not* to work, then scientists accept the evidence that it doesn't work; only findings that scientists try their best to disprove but can't disprove pass scientific muster.

"Facilitated communication" (FC, in which a "facilitator" helps someone who is communicatively impaired type messages on a keyboard but does not influence the messages) illustrates this principle. Someone who says that the real task in research on FC is to show that it *does* work, not that it doesn't, is simply wrong. Science doesn't work that way. Someone who actually understands the scientific way of looking at problems knows that the way to show that FC apparently works is to try very hard to show that it doesn't work and to fail. Only if researchers can't disprove FC are they allowed, using the methods and assumptions of science, to assume that FC works. If researchers are successful in showing that FC *doesn't* work, then scientists will conclude that FC is hokum. Those trying to show that FC *does* work are wasting their time and the time of anyone who listens to them. They're wasting time because FC has already been shown resoundingly *not* to work. The suggestion that the real scientific task is to prove that FC works in the face of overwhelming evidence that it doesn't is much like saying that the real task of science is to prove that there is a raccoon at the dinner table when overwhelming evidence indicates that there is not.

Another example of this principle is the claim that cold fusion had been achieved. The real task of scientists was never to show that cold fusion worked or had been achieved. The task of science was to show that it didn't work and wasn't achieved. Only if people failed in every attempt to show that cold fusion did *not* work would we be led by science to conclude that it must have been achieved.

Contingent Versus Noncontingent Statements

The principle of *contingent versus noncontingent statements* is very important to scientists. Scientists usually qualify their statements by specifying contingencies. They might say something will probably happen only *if* or *when* the conditions are right. The idea is that they describe the *conditions* under which something is likely to happen and those under which it isn't. Almost always in education, results have to be called contingent. For example, the claim that rewards *always* work is baloney; the claim that rewards work has to be qualified. Most teachers know and all scientific investigations have found that rewards have their desired effect only under certain circumstances. It's true that children might be either punished or reinforced by presumed rewards, depending on the circumstances and just how the attention or other rewards are given (Kauffman, Pullen, Mostert, & Trent, 2011). Of course, claiming a contingency that can't be disproved, such as "only if you *really* believe," as is sometimes done with FC, isn't enough. Disproof is still the key.

Replication Versus Idiosyncratic Data

Another misconception about science with special relevance for education is *replication versus idiosyncratic data*. The finding of a single scientist or lab isn't at all convincing to actual scientists. Real scientists are not satisfied unless other people working in other labs can replicate a finding. This was one of the big problems with the cold fusion claim: Other people couldn't make it happen in their labs; only the researchers who claimed they produced it, only those who didn't doubt it, could do it.

Education is especially susceptible to claims that can't be replicated. Finding an effect that can be replicated by other teachers in other locations is particularly important. A single study means relatively little unless it was extraordinarily large and well designed. Even then, definitive evidence can be had only by replication. One of the reasons the programs known as DI are scientifically sound is that the curricula and instructional methods are replicable, and replication has confirmed DI's superior effectiveness (Carnine, Silbert, Kame'enui, & Tarver, 2010).

Observation, Measurement, Reason, and Experiment Versus Philosophy or Ideology

Many people seem not to understand that scientists are concerned about the use of reason or rationality as well as observation and measurement. Actually, scientists are interested in this contrast or competition: *observation, measurement, reason, and experiment versus philosophy or ideology*. Too many educators pride themselves most in their philosophy or ideology and take too little pride in the four contrasting demands of science. Consider at this point just the matter of reason or logic. A science of education requires logic. It requires *more* than logic, but it can't *ignore* logic. Experiment is critical, but so is logical analysis of problems (Engelmann & Carnine, 2011; Engelmann, Bateman, & Lloyd, 2007). Remember that data do not speak for themselves, but when scientists speak they must make sense. That is, they have to be rational.

I recently found an excellent example of *illogic* in reading about standards-based Individual Education Programs (IEPs) in a publication of the National Association of State Directors of Special Education, in which Ahearn (2006) quotes a professor of education: "We must understand that 'ready means never.' If we wait until students are ready to work on challenging standards by virtue of having mastered basic skills, they will never work on challenging standards" (p. 12).

Is this true *only* for teaching children with disabilities, or is it a generalizable principle that we could apply to other problems of education? Think about the training of athletes, musicians, scientists, and, in fact, training in anything in which there are prerequisites for working on more challenging tasks. Ask yourself some reasonable questions. Would you say that basketball players don't really need basic skills in passing, dribbling, shooting, and so on before they play in competition and that judging them not ready for competition because they haven't mastered the basic skills means they will never play in competition? Would you suggest that beginning piano players ought to tackle difficult pieces of music first so that they don't waste time on basic music skills, because if they must first master basic skills then they'll never work on challenging pieces?

Clearly, there is a serious disconnect between ordinary logic applied to other problems of teaching or learning and Ahearn's quotation. Perhaps instructing children with disabilities is a unique case, in that the acquisition of fundamental skills is not necessary for acquiring more advanced skills. But I doubt it. Or maybe the professor of education Ahearn was quoting meant to say that some students are often not appropriately challenged. That assertion may be true. For some students, fundamental skills are a challenge, and some students aren't challenged by fundamentals. But saying that some students are not appropriately challenged is quite different from stating that if students are required to master basic skills before attempting more challenging tasks, then they'll never be asked to take on challenging (i.e., advanced) tasks or standards.

You might also wonder whether the professor of education Ahearn quoted is in la-la land and actually believes that all students can learn whatever is challenging for the majority of students, regardless of what they've mastered previously. In any case, the statement reflects outrageously poor, illogical thinking or careless language or both. I repeat: Science requires more than logical thinking, but it does require logical thinking. There is no illogical science of anything. And when it comes to education, prior learning is the single most important factor to consider in what a student should be expected to learn next. The statement that Ahearn quotes suggests that students with disabilities who are working on tasks that are challenging for *them* don't need to master more basic skills before working on their challenges. I can only hope that someone intended to say that students are always ready to learn their next challenging task only after they've learned more basic skills. But that is not what the statement says.

Gradual Change Versus Paradigm Shift

A lot of education reformers seem to misapprehend the issue of *gradual change versus paradigm shift*. Paradigms do not change often in science, and they are not changed simply by demand, assertion, or act of will. In science, paradigms are changed by data that can't be explained by an older paradigm. And a new paradigm does not necessarily invalidate an old one but might just add to it. For example, quantum mechanics adds to certain aspects of subatomic physics, but it does not overturn or invalidate or replace Newtonian mechanics for macroscopic objects. Many education reformers are particularly fond of the “break the mold” or “breakthrough” idea of educational reform. However, in education, just as in other scientific endeavors, actual paradigm shifts are *extremely* rare. Gradually accumulated evidence is more likely to be a reliable guide to good teaching than is something paradigmatically different from anything we already know.

Theory Versus Fact

Lots of people misunderstand the scientific meaning of *theory versus fact*. In science, a theory is a way of making sense of facts. To a scientist, a theory is not just a guess. A theory is something that for a scientist organizes facts and helps the scientist predict phenomena. In a science of education, theories should help us make sense of research data. “Theory” in education must come to mean what it means in better established sciences. It can’t be a euphemism for ideology or mere guesswork.

PREPARING TEACHERS TO USE A SCIENCE OF EDUCATION

The preparation of teachers has been a highly controversial issue for a very long time. And preparing teachers to put a science of education into practice is just one more controversial aspect of it. One obvious fact about preparing teachers to use a science of education is that we have to have a science of education for them to be prepared to use! Everyone wants better teachers, so that is not the issue. The issue is how to prepare teachers better. If you ask people who don’t know much about teaching what we should do, they are likely to say something like, “Well, get smarter teachers” or “We need teachers who know their subjects, and that’s more important than the kind of teacher training they get.”

How to Train Teachers

How should prospective teachers be selected and trained? That is not an easy question to answer unless you are going to just repeat the same tired old nonsense we’ve been hearing from education reformers for more than 50 years. Trying to answer that question requires some actual knowledge of teaching and schools. If the answer were really simple, either those trying to answer it are too dimwitted to figure it out or there is some sort of conspiracy to keep education from being what it should be.

What are the essential personal characteristics of good teachers? Just how smart does a teacher have to be? Are smarter people better teachers, or at what point does intelligence become irrelevant because just being smarter doesn't make a person a better teacher? What are the essential skills teachers need to be successful? What role does knowledge of each of the following play in making a good teacher: (a) subject matter to be taught, (b) child development, (c) pedagogy or instruction, and (d) behavior management? What other areas of knowledge or expertise are required? To what extent can teachers be prepared before they enter a classroom, and how much (and what) do they simply have to learn on the job? How can we distinguish better teachers from those not as good; that is, how should we rank teachers for reward or recognition and identify those who should be fired for their incompetence or, at least, be told they'd better improve dramatically if they want to keep their jobs? These are not trivial questions. They go to the heart of what teacher preparation is and to the root of controversies about teachers and teaching.

Lots of assumptions and ideologies are related to these questions, but not lots of good thinking and not lots of scientific evidence. Aside from a few obvious characteristics such as not being abusive to students, being fairly intelligent, being reasonably sensitive to the needs of others, and having a relatively high level of energy, we just don't know much about what kind of person makes a better teacher. Aside from the logical assumption that a person can't teach something he or she doesn't know, we are in the dark about how important subject knowledge is. Clearly, people can be failures at teaching what they *do* know. So just knowing something isn't all that's important; knowing how to *teach* it is important if someone is going to be a successful teacher. Teachers who know their subject could be taught to use DI, but that has long been neglected by the education world (Engelmann, 2007).

One thing we can do if we want to make teacher preparation more scientific is look at other types of work that are essentially applied sciences to see how they have made use of the scientific method and put science into practice (Carnine, 2000). It might be impossible to find another profession in which this has been done completely or flawlessly, but that is not essential. We do know that some other professions are way ahead of the teaching profession in making use of science and getting their practitioners to be more consistent in using the practices that science tells them are more effective than just going with their intuition or preferences or some other seat-of-the-pants way of deciding what to do and how to do it.

For example, piloting airplanes and performing surgery are manualized in many ways. Many professions give their trainees manuals because complicated work is involved. The basic idea of a manual is that other people have done this complicated task before and found out how to do it without making a mess of things or creating a disaster. In fact, the manual usually tells trainees as well as experienced professionals how to do something safely, if not best. The reason for following a manual is that responsible professionals do not want to make a fatal error or do something that creates a crisis or unnecessary risk. The manual explains how to avoid a crisis, how to avoid risking disaster, how to do something so that success is more likely than failure. This is why we want the people we entrust with our lives or our health to follow the manual—we want them to do it, as we say, "by the book."

A good manual gives step-by-step instructions based on scientific knowledge and field tests. It is based primarily not on a philosophy or guesswork but rather on what science and logic and experience recommend. A good manual tells us not only how to do something but how to solve problems—how to troubleshoot if something doesn't go right. Why is it taking us so long to manualize the profession of education?

Another way of avoiding disasters that we ought to adapt for education is the checklist. In *The Checklist Manifesto: How to Get Things Right*, Gawande (2009) describes the value of checklists for things like flying airplanes and building skyscrapers in addition to performing surgery. In fact, he suggests, anything that is very complicated can be done far more safely with a really good checklist. The checklist has to be short, focused on the most important things that science and experience tell us, and useful for practitioners. Some people in every profession resist using checklists, but virtually no one receiving professional services thinks that professionals can do without them. Passengers want *their* pilot to use a checklist. Physicians having surgery want *their* surgeon to use a checklist. Using a good checklist is just a way of avoiding an unnecessary calamity. Why haven't checklists become an important part of teaching and preparing teachers?

Perhaps the reason is that education often is not based on scientific information and field tests, as Engelmann (2007) points out so painfully. Educators can't seem to develop a consensus about lots of things, like what they believe children should be able to do and how best to get them to do it. What educators seem to want to do is argue philosophy, not solve problems in a scientific manner. We ought to be aware of what other professions do; for the most part, they apply science, they prepare step-by-step manuals based on practice, they use checklists to help practitioners remember important things, and the more complex the task they undertake the more they see the need for manuals and checklists. In education, we simply don't need to mislead teachers into thinking they can just "wing it" in the classroom.

Suppose we are going to get serious about using manuals and checklists in training teachers. What do we need to know about how a checklist works? Here are some things to remember: (a) A checklist isn't any help if you don't have a specific outcome in mind; (b) you have to know whether what you check off has been done; (c) a checklist does not mean you can be competent without being artful in practicing your profession; (d) a checklist has to be short and designed to avoid common, serious mistakes; and (e) a good checklist does not concentrate power in a particular person, but increases communication and helps people function better as a team.

Now, with all the advantages of checklists, why do so many people despise them, especially for teaching? Well, checklists require close attention to what we're doing, they may make us feel regimented, they point out human frailty—and we like to see ourselves as creative people who are able to improvise and don't need checklists.

Someone could ask, very reasonably, whether we have any manuals and checklists in education. I think we do. For example, much of DI is pretty well manualized, and it has been demonstrated to be a highly effective way of teaching arithmetic and reading, especially to students who are low performing or at risk of academic failure. But we need to develop manuals and checklists for teachers in many more areas of their work. Figure 3 is an example of a *possible* checklist derived from a behavior management text (Kauffman et al., 2011). It may have serious flaws. Before it could be judged sound—reliable and useful—it would need to be field tested like all other checklists that pass muster.

When giving instructions, have I:

- ☐ Made the instruction as simple and clear as possible?
- ☐ Given the instruction in a clear, firm, but polite way?
- ☐ Obtained students' attention before giving the instruction?
- ☐ Given only essential instructions?
- ☐ Given one instruction at a time?
- ☐ Waited a reasonable time for compliance?
- ☐ Monitored compliance?
- ☐ Provided appropriate positive consequences for compliance?

Figure 3. Possible checklist for giving instructions (Kauffman et al., 2011).

CONCLUSION

A science of teacher education is difficult but possible. It first requires a science of education. Such a science requires recognizing impossibilities and inevitabilities, understanding what science is and isn't, and devising manuals and checklists. We must get on with the task of creating useful manuals and checklists for our work. These must be based on reason, field tests, and scientific evidence of effectiveness.

REFERENCES

Ahearn, E. (2006, May). *Standards-based IEPs: Implementation in selected states*. Alexandria, VA: Project Forum, National Association of State Directors of Special Education. Retrieved January 16, 2009, from http://www.projectforum.org/docs/Standards-BasedIEPs-ImplementationinSelectedStates.pdf

Baldwin, J. D. (2008). *Ending the science wars*. Boulder, CO: Paradigm Publishers.

Berliner, D. C. (2002). Educational research: The hardest science of all. *Educational Researcher, 31*(8), 18–20.

Carnine, D. (2000). *Why education experts resist effective practices (and what it would take to make education more like medicine)*. Washington, DC: Fordham Foundation.

Carnine, D. W., Silbert, J., Kame'enui, E. J., & Tarver, S. G. (2010). *Direct instruction reading* (5th ed.). Upper Saddle River, NJ: Merrill Prentice Hall.

Engelmann, S. (2007). *Teaching needy kids in our backward system: 42 years of trying*. Eugene, OR: ADI Press.

Engelmann, S., Bateman, B. D., & Lloyd, J. W. (2007). *Educational logic and illogic*. Eugene, OR: Association for Direct Instruction.

Engelmann, S., & Carnine, D. (2011). *Could John Stuart Mill have saved our schools?* Verona, WI: Attainment.

Gawande, A. (2009). *The checklist manifesto: How to get things right*. New York: Metropolitan Books.

Gould, S. J. (1996a). *Full house: The spread of excellence from Plato to Darwin*. New York: Three Rivers Press.

Gould, S. J. (1996b). *The mismeasure of man* (revised & expanded ed.). New York: Norton.

Kauffman, J. M. (2003). Appearances, stigma, and prevention. *Remedial and Special Education, 24,* 195–198.

Kauffman, J. M. (2005). Waving to Ray Charles: Missing the meaning of disability. *Phi Delta Kappan, 86,* 520–521, 524.

Kauffman, J. M. (2010). *The tragicomedy of public education: Laughing, crying, thinking, fixing*. Verona, WI: Attainment.

Kauffman, J. M. (2011). *Toward a science of education: The battle between rogue and real science*. Verona, WI: Attainment.

Kauffman, J. M., & Lloyd, J. W. (2011). Statistics, data, and special education decisions: Basic links to realities. In J. M. Kauffman & D. P. Hallahan (Eds.), *Handbook of special education* (pp. 27–36). New York: Routledge.

Kauffman, J. M., Pullen, P. L., Mostert, M. P., & Trent, S. C. (2011). *Managing classroom behavior: A reflective case-based approach* (5th ed.) Upper Saddle River, NJ: Pearson.

Landrum, T. J., & Tankersley, M. T. (2004). Science in the schoolhouse: An uninvited guest. *Journal of Learning Disabilities, 37*, 207–212.

Sagan, C. (1996). *The demon-haunted world: Science as a candle in the dark.* New York: Random House.

Snider, V. E. (2006). *Myths and misconceptions about teaching: What really happens in the classroom*. Lanham, MD: Rowman & Littlefield.

Specter, M. (2009). *Denialism: How irrational thinking hinders scientific progress, harms the planet, and threatens our lives*. New York: Penguin.

The Washington Post. (2007, September 10). Save school standards: Congress should resist attempts to water down the No Child Left Behind law, p. A14.