Changing Our Landscape of Inquiry for a New Science of Education

GARY THOMAS
University of Birmingham, U.K.

In this essay, Gary Thomas argues that education research repeatedly makes a mistake first noted by Dewey: it misunderstands our science. This misunderstanding has led to attempts to import various putatively scientific precepts into education inquiry. But in reality, he argues, those "scientific" precepts do not characterize scientific endeavor, which is fluid and plural: science flexes to any angle to answer the questions that are posed in any field. Questions in education concern worlds of practice and social relations where change and corrigibility draw the parameters for inquiry. Education research becomes valuable only when it takes account of the reality of the educational endeavor. Thomas urges us to strive to forge a new science of education based on singular and shared understandings of such practice.

As the twentieth century turned to the twenty-first, calls for more "scientific" research into education increased in number and volume (see, e.g., Levin & O'Donnell, 1999; Mosteller & Boruch, 2002; Slavin, 2002). This phenomenon was, and is, an extension of a trend observed by Lagemann (2000) to have threaded through the twentieth century. In discussing the "elusive science" of education, she notes,

> The most powerful forces to have shaped educational scholarship over the last century have tended to push the field in unfortunate directions—away from closer interactions with policy and practice and toward excessive quantification and scientism. (p. xi)

The trend has not waned in the intervening years; there has been no letup in the call for more scientific research, often on the basis of assumptions about what is scientific.

Underlying these calls for more scientific research are presumptions about the lack of reliable evidence regarding efficacy in education systems and practices. In other words, there is thought to be a dearth of information on what works. The taken-for-granted veracity of these assumptions has impelled

Harvard Educational Review Vol. 82 No. 1 Spring 2012

Western administrations to examine reasons for the "failure" of educational research. In response to such assumptions from its own administration, the National Academy of Sciences formed the Committee on Scientific Principles for Education Research to examine the status of scientific research in education in the United States (Shavelson & Towne, 2002). The Committee concurred with Kaestle's (1993) premise that education research has an "awful reputation" (p. 20):

> The charge to the committee reflects the widespread perception that research in education has not produced the kind of cumulative knowledge garnered from other scientific endeavors. Perhaps even more unflattering, a related indictment leveled at the education research enterprise is that it does not generate knowledge that can inform education practice and policy. The prevailing view is that findings from education research studies are of low quality and are endlessly contested. (Shavelson & Towne, 2002, p. 28)

The consequences of this discourse have included policies such as those embodied in the No Child Left Behind Act (NCLB), which require teachers to use scientifically proven practices in their classrooms. The U.S. Department of Education's (2004) *Toolkit for Teachers* is replete with guidance and instruction on this requirement.

These concerns have not been limited to the United States. Not long before the Committee's deliberations, a similar debate had begun in Europe. The body responsible for assessing the quality of education in the United Kingdom, the Office for Standards in Education, began from premises closely akin to those of the Committee and likewise commissioned a review, which emerged with conclusions highly critical of educational research (Tooley & Darby, 1998). In a preface to the report, the then–chief inspector of schools asserted that "much [educational research] that is published is, on this analysis, at best no more than an irrelevance and a distraction" (p. 5).

While the Committee ultimately emerged with an eclectic view about the surface forms that research can take, there was, underlying its published deliberations, a bedrock of assumptions about the ways science operates that were, it suggested, to be emulated. The Committee called for "rigorous evaluations of innovations" (Shavelson & Towne, 2002, p. 12)—evaluations that would incorporate various "scientific principles" into education research, including the linking of research with theory and the need to replicate and to generalize across studies. It suggested that "science progresses both by advancing new theories or hypotheses and by eliminating theories, hypotheses, or previously accepted facts that have been refuted by newly acquired evidence judged to be definitive" (p. 18). It proceeded: "At a general level, the sciences share a great deal in common, a set of what might be called epistemological or fundamental principles that guide the scientific enterprise" (pp. 50–51). The consequence of all of this has been in the more determined pursuit across the spectrum of education research of what is taken to be scientific (see, e.g., Odom et al., 2005).

The Committee's report represents a particular school of thought about science not uncommon in academic and policy circles. The argument I make in this article is that the view taken there about scientific inquiry, its critical discourse about science, and its selection of cases used to exemplify scientific method provides a narrow view of science and the modus operandi of scientists. The consequence, if such a position is uncontested, will be a terrain of understanding and discourse about inquiry that produces, at best, less useful knowledge than would otherwise be produced and, at worst, inappropriate forms of education inquiry and an impoverished landscape of practice.

In a "folk-view" of science, the stress is placed on what is assumed to be scientific *process*. This emphasis diverts attention from more purposeful attempts to understand the connection found in all scientific inquiry between questions and methods. It seeks first and foremost to *be scientific* about our subject, education, with too scant an attention to the texture of our subject and its susceptibility to various forms of systematic inquiry.

My case is that we cannot talk about scientific method being licensed by the rigorous use of certain routines and techniques, by procedural ground rules, or even by epistemological premises. Each procedural domain in every science is highly peculiar, depending on its subject's form and texture. From astronomy to physics to zoology, the armamentarium of the inquirer is unique. Turney (2004) makes the point well in his discussion of Haack's (2003) commentary on the work of scientists.

> Many who think about science for long probably settle somewhere near the view that it is closely akin to other forms of inquiry, albeit with a much more powerful toolkit . . . there is nothing special about science, no distinctive scientific method. All empirical inquiry, whether undertaken by a detective, an investigative journalist, an anthropologist in the field or a physicist in the lab, demands the same epistemic virtues. Look for evidence as hard as you can, judge carefully what it is worth, and pay scrupulous attention to what it tells you. Then add this evidence to existing experience and use it to reason out your best-informed conjectures. (paras. 1, 3, 4)

Scientific research is, in other words, fluid and multifarious. Assuming otherwise—taking an unyieldingly monistic view about the procedures characterizing the enterprise of inquiry—leads us on some epistemological wild goose chases.

This point is made commonly in the sociology and philosophy of science. Becker (1998), for example, notes the work of researchers such as Latour and Woolgar (1979) and Lynch (1985), who revealed how natural scientists work "in ways never mentioned in their formal statements of method, hiding 'shop floor practice'" (Becker, 1998, p. 5). Or, as Medawar (1982) suggests of his own discipline, biology, the special methods and procedures that are supposed to be used by scientists represent merely "the postures we choose to be seen in when the curtain goes up and the public sees us" (p. 88).

Their message is that the notion of a scientific method is a pretense. As Albert Einstein asserted, there is no correct method; the creative scientist must be an "unscrupulous opportunist."[1] The essence of science, he said, is the seeking "in whatever manner is suitable, a simplified and lucid image of the world . . . There is no logical path, but only intuition" (cited in Holton, 1995, p. 168). He explained that "the whole of science is nothing more than a refinement of everyday thinking" (Einstein, 1936, p. 351). There is, in other words, no privileged set of methods in science quarantined away from everyday thinking. We are wrong if we assume that there is some archetypal set of procedures that can collectively be laid over any focus of inquiry that will make our deliberations scientific.

Lest anyone get the wrong idea, I should note that in writing this essay, my thesis is emphatically not an antiscience one. Indeed, I share Isaiah Berlin's (1979) view that the sciences' espousal and use of rationalism probably marks the major achievement of the human mind. My case is that there is no core to scientific method, no charmed circle of precepts and processes that lead the incipiently scientific inquirer to the sunlit uplands of scientific inquiry. My argument is about the ways that we choose to be scientific in education inquiry and the consequences that such choices have for the nature and growth of our field of endeavor—our own science. Before looking at those choices, I examine the key assumptions made about science in the current discourse. The first of these is about the cumulative nature of scientific knowledge.

Science and Cumulation

The Committee on Scientific Principles for Education Research placed great stress on the ability of science to accumulate knowledge. Indeed, a whole chapter of its exegesis was devoted to this subject.

> Theory and method build on one another both as a contributor to and a consequence of empirical observations and assertions about knowledge. New knowledge gained from increased precision in measurement (say) increases the accuracy of theory. An increasingly accurate theory suggests the possibility of new measurement techniques. Application of these new measurement techniques, in turn, produces new empirical evidence, and so the cycle continues. (Shavelson & Towne, 2002, p. 48)

Philosopher of science George Canguilhem (1994) suggests that scientific advance cannot be defined much more satisfactorily than the "elimination of the false by the true" (p. 41). The seeking of patterns—of generalizing theory—has in natural science enabled that eventual elimination of false by true via cumulation. Or at least (for those who balk at the starkness of false versus true) it has enabled an elimination of less reliable knowledge in favor of more reliable knowledge (Ziman, 1991).

Education has not produced such cumulation. But this is *not* because we haven't got the principles behind cumulation quite right. In education, the process of elimination of less reliable knowledge by more reliable knowledge happens at a different epistemic stratum, for the knowledge we educators trade in is different from the knowledge of the physicist or plant scientist. It is discovered at an individual and local level, as I shall explore in more detail later in this essay.

Today's student of aeronautics undeniably knows more about powered flight than did the Wright brothers, and even an average high school student today almost certainly knows more about the crude facts of the circulation of blood than William Harvey did in 1628. The knowledge in these fields has been uncomplicatedly cumulative, and we benefit from the accumulation. But is it right to imply that teachers today do not know more about teaching than their collegial ancestors? My own view is that teachers do indeed know more today—much more—but we know it from a cultivated accumulation of understandings rather than an accumulation of facts. We know it from the development of a community of inquiry that is continually imagining and assaying, conjecturing and refuting, reviewing and reassessing personal and collective experience. The improvement is attributable to the kind of knowledge from which educators operate: knowledge coming from immersion and informed reflection. This is the cumulation of our science.

It was philosopher Gilbert Ryle (1990) who drew attention to the distinction between *know-that* knowledge and *know-how* knowledge. The know-how knowledge of the educator is tacit, practical knowledge, and the practicing teacher's know-how knowledge may or may not be better than that of a colleague from a previous generation, or indeed that of Froebel or Pestalozzi. Whether it is better will depend on the quality of the practical reflection, the use of personal experience, and the weighing of others' experience, and it is in the process of guided use of experience and applied inquiry that the accumulated quality of the developing teacher's practice is assured. Know-that knowledge is an accreted knowledge of facts: collectable, cumulative, progressive, and clearly demonstrable in the natural sciences. But any expectation about cumulation in our scientific inquiry in education has to rest on an accumulation not of generalizable facts but of *understandings* drawn from and assessed in the context of one's own experiences and the experiences of others. It rests, in other words, in the cultivation of provisional, tentative models for interpretation and analysis.

Science and Generalization

Generalization, replication, and prediction are taken to be key elements of scientific method, and the Committee again provided a clear statement of received wisdom here. Its members averred that one of the fundamental principles that guide the scientific enterprise is the use of "observational methods linked to theory that enable other scientists to verify their accuracy, and recognizing the importance of both independent replication and generalization," going on to say that "the long-term goal of much of science is to produce theory that can offer a stable encapsulation of 'facts' that generalizes beyond the particular" (Shavelson & Towne, 2002, p. 51).

The key element here is generalization, for replication and prediction are merely underlings to this executive. Generalization is taken in this view of science to be the sine qua non—is assumed to lie at the root of the development of theory and thus explanation and prediction. With generalizing theory applied retrospectively, one can explain; applying it prospectively, one can predict—or so the reasoning goes.

It is the construction of *generalized* and *generalizable* knowledge of a particular quality that is taken to be the cynosure of the scientist—and therefore of the social scientist and, more particularly, the education researcher. Indeed, it is in generality or universals that we find claims about the distinctive offer of social science—distinctive, that is, over and above the common insight of the layperson. We seek to capture regularity, and its encapsulation often goes under the title of *theory*. The nature of the regularity putatively discovered provides the warrant for the claims of the research. If our generalizations were no better than those of the layperson, what would be the point of calling our endeavors social *science*?

From experiment to ethnography, generalizing power holds great appeal for social scientists. In experiment, the use of sampling and inferential statistics concerns the degree to which general conclusions can be drawn from particular circumstances (see Campbell, 1957, for a classic discussion). The same can be true for ethnography: Becker (1998) talks of ethnography's "generalizing tricks" (p. 3). Nadel (1957) discusses theory as consisting of interconnected generalizations existing in such a way that "observable consequences logically follow" (p. 1).

I have two concerns with some of these assumptions concerning generalization. The first is that generalizing theory, while possible in certain natural sciences, is possible only to a far more limited extent in others. Particularly in the social sciences, it is problematic. The second is that, even in natural science, while generalization is essential much of the time, it is not always possible, nor is it always to the point. The rational endeavor of science is not always dependent on generalization, and we should restrain a first impulse to make abstract, to generalize, to find principles, to synthesize, and then to call all of this "theory" and to engage in a pretense that reliable explanation and prediction are possible because the theory owes its origins to a notionally scientific method. The set of assumptions behind this process constrains our capacity to examine and understand the individual—the idiographic—and may therein distort our ability to see and understand clearly the essence of our subject.

In education, there are special reasons for being circumspect about generalization's offer. The undeniably large number of factors that come into play in education should make us instinctively cautious about what might be called the *generalizing scientific endeavor*. Again, Einstein (1941) offers a relevant insight:

> When the number of factors coming into play in a phenomenological complex is too large, scientific method in most cases fails us. One need only think of the weather, in which case prediction even for a few days ahead is impossible. (para. 8)

Prediction based on generalization is sometimes easy; sometimes it is hard. Whether it's easy or hard depends on Einstein's "number of factors coming into play in a phenomenological complex." In his now-classic critique of social science, Alasdair MacIntyre (1985) puts the case rather more fully. He notes that social science's generalizations are unsatisfactory as scientific generalizations (in the "natural" scientific sense) because social scientists cannot specify the conditions under which they operate. Social science has made little or no progress, asserts MacIntyre, because it persistently tries to employ a misunderstanding of what natural scientists do in employing generalization.

MacIntyre's conclusions are valid not only for social science but for many forms of inquiry. I can predict with some certainty the rise of the sun tomorrow. I can predict that if I apply a load to a spring that the spring will stretch according to Hooke's Law. However, Einstein's point about the weather and the phenomenological complex is only too clear to me, living in a meteorologically unstable archipelago in the northern Atlantic Ocean (that is to say, the British Isles). Neither I, nor the hugely sophisticated modeling infrastructure of my country's Meteorological Office (complete with second-to-second satellite readouts and a staff of thousands) can predict what the weather will be outside my front door three or four days from now. In fact, my guess is almost as good as the Meteorological Office's (see Rescher, 1998).[2]

It's a mixed bag with prediction. Sometimes it's easy and accurate. Sometimes it's difficult and wrong. And the issue for educators—and social scientists more generally—is that we need to be aware of the parameters that define the ease and difficulty of prediction. More importantly, we need to think about whether prediction—about, for example, what works—is the essence of our game.

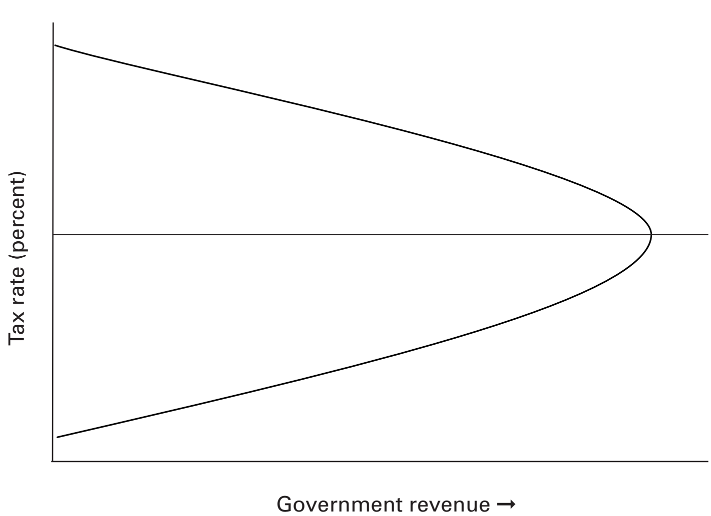
Put differently, we need to be aware of the kind of science in which we are engaging. Is it first and foremost the kind of science that seeks generalizations from which predictions can be made, generalizations depending for their legitimacy on carefully controlled trials? Is this the nature of the education research enterprise? Pursuing that question, let's examine another area of inquiry where Einstein's "phenomenological complex" is broad: economics. Economics is not dissimilar to education: the factors at play in any inquiry interrelate erratically and with high degrees of seeming randomness; agency is untethered, and its role is major. For these reasons, and many more, prediction is difficult, even in the very short term.

Economists emerged prior to the post-2007 recession with some highly lauded modeling about the future of the global economy based on portfolio theory and asset-pricing models. Their message was, in short, that there was no need for concern about the economic future. The models they developed were highly complex and furnished explanations about why the growth of the global economy would continue unabated.

But they turned out to be wrong—very wrong. As mathematician and social scientist N. N. Taleb (2008) has pointed out, Nobel Prize–winning economists failed to foresee the onset of the most serious recession since the 1930s. He pinpoints the culpability for this failure in the near-universal reverence paid to model-making and theory-building in the social sciences. In his exegesis he rejects the seeking of pattern where there are no patterns (or, at least, only perpetually shifting patterns). Model making and theory building, built on tidying generalizations and a fondness for all things quantitative in the social sciences, are chimerical, suggests Taleb. Events have proved him correct. There is very little useful prediction to be had, and the "theory" that has materialized out of the methodological predilections of many social scientists, particularly these economists, has been profoundly misleading.

The celebrated mathematician Martin Gardner (1996) made a very similar point about the genesis of inappropriately tidy modeling. The Laffer curve (see figure 1) was a construct devised by supply-side economists to explain the relationship of taxation levels to government income. There was an optimal point for government revenues, the theory predicted, between taxation level and personal income, since if taxation were at 100 percent, no one would have any incentive to work, while at zero, there would be high incentive yet no tax income. The best point for the government must lie somewhere in between,

FIGURE 1 *The Laffer Curve*



the curve predicted, and the concept was used to argue for lower taxation rates producing higher revenue for government.

Gardner parodied the naïveté of the modeling with what he called the "neo-Laffer curve" (see figure 2). The neo-Laffer curve is based on actual data for the U.S. economy over fifty years, which discloses what Gardner (1996) calls "a swarm of densely packed points" (p. 132). Plotting from these data produces a "technosnarl" wherein it is impossible to predict at what point tax rate will maximize government revenue. We can, in other words, collect accurate data, but these will not knit together in the way that we can reliably expect of data describing a spring's extension. It is subject to all the vagaries of the social world that MacIntyre described.

Education is no better on prediction than are meteorology or economics; nor should we expect it to be. We cannot say very accurately from research what will be a good way of teaching—what will have good results and what will have poor results—even after expensive evaluations. But the problem is *not* that we are not doing this kind of research well enough. The problem is that we are trying to do it at all. This is indeed the lesson that seems to be emerging ten years after NCLB, when teachers and administrators should, by legally backed instruction, have had their practices and policies determined solely by scientifically based research. If teachers and administrators have been obeying instructions—applying "rigorous, systematic and objective procedures to obtain reliable and valid knowledge relevant to educational activities and progress" (U.S. Department of Education, 2004, p. 30)—there is no evidence of ensuing benefit in the PISA (OECD/UNESCO, 2003; OECD, 2010) statistics for the performance of American students for the period since that injunction (see table 1).

FIGURE 2 *The Neo-Laffer Curve*

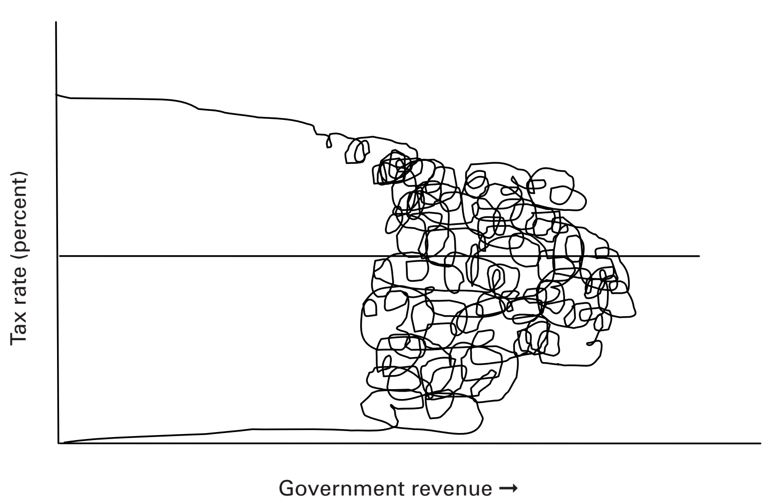


TABLE 1 *PISA test scores for 15 year olds in the United States, pre- and post-NCLB*

Subject	*Year*	
	2000	*2009*
Reading	504	500
Mathematics	493	487
Science	499	502

My argument is that the lack of improvement should not be surprising. Teaching is neither open to nor penetrable by such supposedly scientific research. Certain forms of inquiry seem to be endemically prone to fail to provide reliable information of the hoped-for kind on certain topics. Across the social sciences, there is a notable and persistent lack of success in producing cumulative knowledge and progressive theory; there is unrelenting frustration in the attempt to link generalizing theory and practice (see Thomas, 1997, 2007).

The answer to this failure lies not in formulas-based attempts to do better and more systematically what has already disappointed. Rather, it lies in a shift in the way that inquiry into our subject is thought about; a shift in the way that the education research enterprise—our science—is conceived. None of this is to imply that attempts to be scientific are misplaced. Far from it. Instead, I suggest that science takes varied forms—it flexes to any angle to answer the questions that are posed in any field—and the form that is appropriate for inquiry in education will be very different from others, just as each scientific inquiry form and process is discrete. The landscape of inquiry, to which I shall return later, needs to be less concerned with procedure and prediction and more concerned with the present, its vagaries and its interstices. As Thomas Carlyle (1829) put it many years ago, "Our grand business undoubtedly is, not to see what lies dimly at a distance, but to *do* what lies clearly at hand." For educators and their worlds, it is the culture of the present, a culture of *doing*, of practice, that needs to form our laboratory.

Scientific Research and the Individual Case: Playing with Ideas

In one of his lectures, Schön (1987) made the following point:

> There is . . . the notion that the more general and the more theoretical the knowledge, the higher it is. I remember once being quite recently at a school of education, and a graduate student was in a seminar that I was doing, and she was working with nurses, and she said something I thought was interesting. And I asked her if she would give me an example. And she then gave me a proposition which was just as general as the first proposition. So I asked again for an

> example, and she gave me a proposition which was just slightly less general. And I asked again, and I finally got an example. And I asked her afterwards if she thought it was strange that it took three or four tries to get an example, and she said she *did* think it was strange, and she didn't understand why she'd done that. And I think it is because she had been socialized to an institution where, tacitly and automatically, we believe that the only thing that really counts and the only thing that's really of value is *theory*, and the higher and the more abstract and the more general the theory, the higher the status it is. Under such conditions it's very difficult to give more or less concrete examples. (para. 5)

Dewey (1920) expressed the same kind of concern about educational inquiry, saying that what was needed was more specific inquiry into a multitude of specific structures and interactions in education. Both Schön and Dewey were suggesting that we should inhibit a first impulse as a community of inquirers to make abstract, to generalize, to find principles, to synthesize, and to predict on the basis of that synthesis. The impulse automatically to make abstract constrains our capacity and curtails our desire to research—to look more deeply at—the individual situation and to reflect on that situation in order to improve practice. In other words, it leads us away from more valuable ways of thinking about the subjects of our interest. As Wittgenstein (1958) warned, the "craving for generality" and the attendant "contemptuous attitude toward the particular case" make us "dismiss as irrelevant the concrete cases which alone could have helped to understand" (pp. 17–20).

Wittgenstein's comments represent a set of counterpositions about inquiry that is, of course, as old as the Athenian hills. The Platonic position, which has a fine and productive record in reductionist science of one kind or another, was that we should look for the *simile in multis*—look, in other words, for the *essence*, the theory, that will capture the truth evinced by many cases. The Aristotelian position, by contrast, was that for many kinds of knowledge we progress only by using our practical reasoning (*phronesis*), craft knowledge (*technê*), or tacit knowing—the stuff we know because of our experience. It's the stuff we learn "on the job," and it is explicable and researchable only in terms of the particular context. It cannot therefore be researched via a reduction to general principles (Thomas, 2011).

It is the latter, contextual, case-based research on which our science should properly focus and on which it should be based. Scientific progress of the simile in multis kind is unequivocally cumulative, but the inquiry canvass of the educator is more concerned with phronesis and technê.

In this, education is not alone. I'd like to offer two illustrative case studies about the individual and idiosyncratic process of science and what I would call its *play*. Progress in these instances and myriad others rests neither on deliberate attempts to follow scientific method nor on adherence to some set of scientific precepts. Far from it. Rather, progress rests on playing with ideas and the development of an explanatory narrative via the explication of the singular. Here, the work of the scientists is similar to that of educators.

The first case study is from astronomy. In 1967, an astronomical phenomenon (to which the name *pulsar* was later given) was discovered by the astronomer Jocelyn Bell Burnell and her colleagues with the then-new radio telescope at Jodrell Bank near Manchester, England. How did she discover it? It started first with noticing. Doing her routine work, Bell Burnell noticed something unusual. Something in a distant galaxy was giving off massive pulses of radio energy at fantastically regular intervals. The rhythm was as regular as an atomic clock. This was a mystery, for nothing else in the known universe did anything like this. After the noticing, it was time for conjecture. At first, Bell Burnell and her colleagues conjectured that this was a message from intelligent life: the pulsar was given the name LGM-1 (Little Green Men 1). There was no better explanation for this regular-as-clockwork pulsing. Bell Burnell and her colleagues had to make judgments based on their existing knowledge of phenomena such as this, and one of their first explanatory forays, only half-jokingly, suggested aliens trying to communicate.

But one conjecture was clearly insufficient. To go beyond this, Bell Burnell and her astronomer colleagues began to look in different ways at the pulse. By using their existing knowledge and imaginative speculation—playing with ideas, putting forward and rejecting or accepting plausible hypotheses—they eventually reasoned that these regular pulses were the flashes of radiation from a collapsed neutron star that was spinning at an incredible speed, giving out a directed and extraordinarily condensed beam of light, like a lighthouse.

The point I am trying to make from this example is that this conclusion was arrived at not through experimental method—or indeed any method. The conclusion was drawn by studying just one or two pulsars in a great deal of detail and with a great deal of thinking. Using her own knowledge, the knowledge of the discipline, and the tools of her trade, Bell Burnell drew hugely significant scientific conclusions from this single case. The conclusions won her team a Nobel Prize.

The second case study is from paleoanthropology. This is how the celebrated paleoanthropologist Louis Leakey described how and why he and his wife, Mary Leakey, started work on hunting for fossils containing evidence about our human forebears in the Olduvai Gorge in Tanzania: "For some reason both of us had been drawn again and again to this particular site" (Leakey & van Lawick, 1963, p. 134).

"*For some reason*" is fascinating. "*For some reason*" says that the reason they started work in this location was beyond their articulable knowledge. It gives a clue about the way that a scientific practitioner tacitly uses evidence—evidence built into something we know as *experience.* For this is what experience is: personal, tacit knowledge built out of information—data, evidence—accumulated both deliberately and fortuitously. Ideas often emerge out of confluences of circumstantial evidence in the minds of those who are steeped in a problem, a practice, a discourse, or a technology, whether we are talking about education, astronomy, or paleoanthropology. Whatever the science, its practitioners swim in the discourse and play with the evidence. Often those scientists have a feeling, a hunch, that this way or that way is the right way to proceed, without being able to articulate its provenance in evidence. There is a playing around with bits and pieces of everyday evidence, bits and pieces of tacit knowledge, that in some way enables practitioners to discover the way to proceed—to conjecture and reason.

Beyond intuiting where to look, the Leakeys played with evidence by collecting fragments of bone and other material and piecing together stories—plausible, intuitively guided accounts—of the way that this prehuman may have led to that one. The evidence they used was not monochromatic. There was knowledge from geology about the age of rocks in which materials were found; there was carbon dating; and there was even knowledge from psychology and physiology. For example, Louis Leakey worked out that one particular prehuman had an ability to use language by looking at the roots of the fossil specimen's teeth: a small depression in the jaw-bone called the canine fossa can be shown to be related to the production of speech, for the little dip makes space for a muscle used in language production (Leakey & van Lawick, 1963). This evidence was acquired through curiosity, serendipity, search, questioning, surprise, and accumulated experience and was melded into a narrative through playing with ideas. It was advanced case study—painstakingly careful examination of the particular. None of this was formally experimental, nor even was there any special method. Yet few would doubt that this case study was science of the most sophisticated kind. From Jocelyn Bell Burnell's pulsar to the Leakeys' fossils, the study was of *cases.*

All scientists—whether physicists, chemists, biologists, paleoanthropologists, or educators—use particular kinds of evidence and play with it in particular ways relevant to their fields of work and the methodological traditions that have developed there. But methods are forged around questions; the methods have been the servants, not the executive directors. It is the questions that are important. The natural sciences gained purchase on the world not principally by discovering better methods or by deliberately going out to link theory and practice. On the contrary, they did it by asking good questions—questions germane to the worlds that they studied—and they contrived procedural routes, protean and specialized, to answer these questions.

If we fail to understand that different sciences take different routes, and we consequently try to superimpose the route of one domain of conspicuously successful scientific inquiry onto another, it is as if we are trying to play different games by the same rules. As Dewey (1923) put it:

> It may conduce to immediate ease or momentary efficiency to seek an answer for questions outside of education, in some material which already has scientific prestige. But such seeking is an abdication, a surrender. In the end it only lessens the chances that education in actual operation will provide the materials for an improved science. (p. 77)

There is nothing wrong with the use of metaphor or analogy for the purposes of illustration, for carrying familiar images and ideas from one domain to another, but the process should not move beyond illustration, Dewey implies. If correspondences are imagined too creatively, too forcefully, extended too far—the consequence may be in our believing something profoundly misleading.

Once analogies have been overdrawn, metaphors overextended, the way is clear for unwelcome sequelae. One example of such overextension is the set of assumptions that accompany the similarity drawn with medical science in its use of randomized controlled trials (RCTs). Slavin (2002), for example, has gone so far as to suggest that "it is the randomized clinical trial—more than any single medical breakthrough—that has transformed medicine" (p. 16).[3] This offers to us—and it has been accepted in many education quarters—imagery painted in fluorescent colors for the success of a research method based on a putative correspondence between medical research and education research. It is worth some deconstructing. Aside from the accuracy and relevance of the correspondence itself (see, e.g., Norman, 2003), can we really say that RCTs are responsible for medicine's progress? Surely it is curiosity, questioning, imagination, insight, and ingenuity that are responsible for the advances in medicine. If we want to understand the armatures and levers of advance, we are unlikely to be rewarded by an analysis of the role of RCTs. Rather, we would do better to focus on the ways that tacit knowledge is used by practitioners and by the ways that intelligent noticing emerges within and outside local communities and infrastructures of inquiry. If one looks at well-documented advances in any science or technology—in the understanding of superconductivity, the discovery of penicillin, the discovery of helicobacter pylori (the bacteria that cause ulcers), the invention of nylon, or hundreds more—it becomes clear that it is inspiration, creativity, and imagination coalescing and crystallizing in extraordinary ways that enable advance (Thagard, 1998). Advances emerge from able people working with their personal knowledge, the tools of their trade, and their immersion in the ideas of their intellectual communities. The elevation of RCTs' role in the "extraordinary advances" Slavin (2002, p. 16) admires in medicine is misplaced, for the RCTs perform only a simple, confirmatory role.

The misplacement resides in a misunderstanding of different kinds of scientific endeavor and advance. In explicating this point, Medawar (1982) distinguishes between the "poetic" and the "bookkeeping" varieties of scientific endeavor. The poetic work is creative and imaginative; it pulses through the center of scientific advance. The bookkeeping work lies in the grinding follow-up, the refinement, the clarification, the confirmation, and the establishment of exceptions. While it's tough to do, you don't have to be a genius to do it. Both kinds of work—poetic and bookkeeping—are essential, but the role and character of each is entirely different. The bookkeeping work could not exist without the poetic: there would be no bookkeeping to do, since the advance in thought would not have been made.

RCTs certainly play a vital role in validating effectiveness in medicine, but they are part of the bookkeeping aspect of science—the mundane and rather dull side that is necessary once the inspirational advances have been made. Certainly, as part of the bookkeeping side of their work, plant scientists, pharmacologists, and medics use RCTs very successfully. However, the great majority of scientists do not use RCTs because they don't need to.[4] Even where they are necessary, we should be aware of their significance: their role is an essential but humdrum one.

Education's Landscape of Inquiry

I should summarize my argument so far. It is that:

- Systematic inquiry in different domains rarely follows the same methodological avenues, nor even does it adhere to the same precepts.
- Not all science is about generalization; it is certainly about explanation, but explanation does not always require generalization from the few to the many.
- The cumulation of know-how knowledge differs from the cumulation of know-that knowledge.
- Education's cumulation is by the cultivation and sharing of practical understandings, not an accumulation of facts.
- The methods used in a scientific endeavor must *emerge from* questions; they should not be prescribed in advance.

I wish now to look in more detail at that last point—at the questions posed in educational research and the types of knowledge in which such questions are situated. A field of inquiry spawns particular kinds of questions, and those questions in education, my argument so far suggests, should exist at the level of practice. Our landscape of inquiry cannot exist with what Flyvbjerg (2001, 2006a, 2006b) calls "Big Questions"[5] because the generalized, depersonalized answers to those questions evaporate almost as soon as they emerge. If one looks, for example, at Ravitch's (2000, 2010) reviews of attempts at education reform over a century, one might suggest, only half-facetiously, that there are two equally valid answers to the question "What works?"—namely, "Everything works" and "Nothing works." Everything and anything works if it is supported with gusto in an environment for which it is designed. But the message to emerge from a generation of evaluation is that the gains don't survive and they don't extrapolate. As Ravitch (2010) puts it:

> If there is one consistent lesson that one gleans by studying school reform over the past century, it is the danger of taking a good idea and expanding it rapidly, spreading it thin. What is stunningly successful in a small setting, nurtured by its

> founders and brought to life by a cadre of passionate teachers, seldom survives the transition when it is turned into a large-scale reform. (p. 146)

Our landscape of inquiry exists not at the level of these big "what works" questions but at the level of personalized questions posed locally. It exists in the dynamic of teachers' work, in everyday judgments; and these are judgments wherein generalizations and theories exist and thrive, certainly, but do so at a different epistemic stratum from those generalizations garnered as part of most natural scientific inquiry.

The substrate of our inquiry lies in practices that are, if nothing else, context-bound and corrigible. Practical knowledge, the everyday practice of teachers' work—and thus inquiry about that knowledge—must take as a starting point variability, transitoriness, and contingency rather than consistency and law-like regularity. Knowledge gained as part of practice is tentative, provisional, and characterized by change; it involves recursions, iterations, revisions, and repeated review. In this, it is no different from the repeated conjecture and refutation of the natural scientist, but it develops personally, locally. The knowledge gained about what works, therefore, has an important predicate: it is "what works *for me*."

In practicing and researching teaching, local considerations will always obtain, and these will rest in what E. D. Hirsch (1976) calls "calculations of probability based on an insider's knowledge" (p. 18). Teachers develop and use what Hirsch (1976) calls "local hermeneutics" in craft knowledge—the rule of thumb and the knack. Local hermeneutics take shape in informal, changeable, evanescent devices for interpretation, analysis, and action. Stanley Fish (1989) suggests that practitioners develop schemata that "vary with the contextual circumstances of an ongoing practice; as those circumstances change, the very meaning of the rule (the instructions it is understood to give) changes too" (p. 317). These schemata, these personal rules, are too fragile to be generalized. Their hallmark is in their malleability, their corrigibility, as distinct from the rigidity that inevitably comes from a set of "what works"–led presuppositions.

So it is out of local hermeneutics—craft knowledge—that the fabric of educational knowledge is woven; it is fashioned out of teachers' tacit knowledge about whether to make eye contact, when to feign crossness, how to respond to an interruption, what kind of question to ask, how to pretend disappointment, what language to employ in developing a theme. All of this knowledge comes from evidence amassed and adduced out of practical experience.

Fish (1989) speaks of "heuristic questions" in practice. A heuristic is a rule of thumb that leads to conclusions and predictions that may not be correct all of the time but are good enough most of the time (see Newell & Shaw, 1957). It is a temporary model developed out of experience and takes as one of its starting points the fact that, in most circumstances, we are unable to garner enough information about all of the relevant variables to warrant the drawing of watertight conclusions about the veracity of any proposition we may make. Such situations are commonplace in the human sciences, and possibly ubiquitous in the research tableaux of education. In the practice of education, teachers have to make judgments about what is important and develop pathways to action based on these judgments. These local calculations, these insider knowledges, are forged in the contextual crucible of practical activity. It is these that surely lie at the heart of inquiry for the teacher—at the heart of our science.[6]

Our Science in Practice

The landscape of our inquiry takes shape in the tacit and explicit forms of problem solving that are part of, and emerge from, our practice and our business of everyday doing and being—our work, our interactions with others, our reading, our discussions. They are characterized by revision and informal experiment, by the flexibility and changeability that hallmark local practice.

Worked into teachers' thinking, these forms of problem solving are the "deliberative pedagogies" spoken of by Cochran-Smith and Lytle (2006, p. 692). Teachers work together feedback from informal trials, ideas, propositions, speculations—meshing together practice and research—to cultivate and develop their professional work. Reid and Valle (2004) make insightful points about the ways that such deliberative pedagogy may actually be built out of, and integrated with, personal, tacit knowledge. They say that teachers need to approach their work as scholar-practitioners, sharpening the tools of critical inquiry: "observation, conferencing, and interviewing; generating anecdotal records; taking and analyzing field notes; constructing sociograms; analyzing student work and portfolios" (p. 474). Research is thus folded into practice. Hennessy, Warwick, and Mercer (2011) describe related work done by researchers and teachers using digital video to critically evaluate their own and others' practice. This was genuinely dialogic co-inquiry wherein, as the authors put it, "collaborative theory-building" happened. Out of the process, pedagogical rationales were shifted and altered. Zeichner (2007) explains how disparate, singular studies of this kind may be aggregated and cumulated to offer to a community of inquiry.

While there should be this development of a community of inquiry, this should not mean that we lose sight of the fact that the focus of the teacher's work is, in the first place, singular and local, as indeed it was in the work of astronomer Jocelyn Bell Burnell and paleoanthropologists Louis and Mary Leakey. It is about explaining through the construction of potential narratives. As Becker (1998) suggests, the notion of a straightforward "cause" is mistaken in social research. His advice is that we should seek narrative rather than cause: "Assume that whatever you want to study has, not causes, but a history, a story, a narrative, a 'first this happened, then that happened, and then the other happened, and it ended up like this'" (pp. 60–61). It is the resilience of such narrative, such rational conjecture, to reasoned attempts at refutation that lies at the center of science in our field of education.

The construction of narrative characterizes all development of theory—or, perhaps better, phronesis—at the practical level, incorporating shifts of view and changes of mind (see Fenstermacher, 1994; Huberman, 1999; Marton & Booth, 1997; Weiss, 1995). I have made it clear elsewhere (Thomas, 2010) that I see "exemplary knowledge" as the key, and I am talking about example viewed and heard potentially in the context of another's experience—another's horizon[7]—but used in the context of one's own, where the horizon changes. The example is not taken to be representative, typical, or standard; nor is it exemplary in the sense of being a model. Rather, it is taken to be a particular representation given in context and understood in that context. It is penetrable only in the context of one's own experience—in the context of one's phronesis, rather than some distal theory.

In practical terms, this leaves us with a science of education concerned with local study conducted by practitioners. Rather than seeking guidance for practice from bodies of theory or generalized knowledge, such study can offer a series of ways of proceeding based in exemplary knowledge. Through such exemplary knowledge, one can make connections between another's experience and one's own, seeing links, having insights, building what Berger and Luckmann (1979) call "multiple realities" (p. 21). A fine example appeared recently in the pages of this journal, wherein Wright (2010) chronicled, reflected on, and analyzed the emotional stasis and eventual thawing and trust of a little girl with whom he was working. Refreshingly free of the code-theoretical or quasi-technical constructs that so often characterize accounts of breakdown in learning or emotional development, Wright's explanation about the girl's withdrawal came directly from what he saw and what he knew. His intuitions about how to behave with her came from his own experiences as a person and a professional, one with experience of other people, one who approaches others with humanity, understanding, and a will to succeed. We read in the context of our own experience, our own horizons of understanding.[8]

Maclure (2009) suggests that we should

> adopt what Brian Massumi calls an "exemplary method" (2002: 17)—in other words, working theory through examples. This does not mean enhancing the street cred of theory by sticking some examples "into" it, which would amount to mere "application." Rather, it would mean exploiting the strange play of the example . . . Examples necessarily involve details . . . their success hinges on them. So conceptual development is worked at the level of singularity and specificity. (p. 6)

She describes how in a project with young children at school, out of a "wearying mass" of ethnographic data (videos and field notes), something would catch her attention, often in a meeting with others, and start to form itself into an example.

> It is hard to describe how this happens, since you cannot recognize an example right at the point of its emergence. One way to describe its beginnings would be as a kind of *glow*: some detail—a fieldnote fragment or video image—starts to glimmer, gathering our attention. (Maclure, 2009, p. 7)

I suspect that this "glow" is similar to the excitement that Bell Burnell felt on gaining her insights about the pulsar's beam, or the thrill that the Leakeys must have experienced on appreciating the significance of each new fossil. It comes from hunching and hunting—having hunches about what might be, hunting for evidence, developing an explanation, and continually testing that explanation in the light of new evidence. This is our science.

The locality of this science is highlighted by the celebrated sociologist C. Wright Mills (1970) in his advice to new social researchers. He said the following about what he called the "intellectual craftsmanship" of the social scientist, and his advice (in the gendered language of his time) is apposite for the educator:

> Be a good craftsman: Avoid any rigid set of procedures . . . Avoid the fetishism of method and technique. Urge the rehabilitation of the unpretentious intellectual craftsman, and try to become such a craftsman yourself. Let every man be his own methodologist; let every man be his own theorist; let theory and method again become part of the practice of craft. Stand for the primacy of the individual scholar; stand opposed to the ascendancy of research teams of technicians. (p. 224)

The stance of Wright Mills's craftsperson resists the detached focus on what might be called the "theorized general." The latter, the disconnected researcher seeking disinterested answers and delivering them to practitioners, has consequences beyond an unrelentingly unproductive set of research findings. It has consequences also for the way that we think about our work as educators—how we reflect about curriculum, learning, and change.

From comparative analysis, Louis (1992, p. 291) gives some pointers to the ways that research and practice may be conjoined and integrated as Wright Mills urged. She observes that the traditional relationship of research to practice in the United States (and, I would add, in the United Kingdom and the English-speaking world more generally) is knowledge driven, with the emphasis placed on the dissemination of researchers' research to teachers rather than on the facilitation of teachers' own practice-based inquiry. Using examples from the Netherlands and Denmark, she notes, though, that there is an alternative where the emphasis is on problem solving rather than dissemination: there are local guidance centers, peer consultancy, high levels of in-service education, school-based action research.

Extending the comparative view to Finland, which is renowned for the consistent success of its education (Simola, 2005), there has been a resolute resistance to what might be called directive research and an equally tenacious insistence on advanced education for teachers, with master's degree achievement necessary for permanent employment (Ojala, 2005; Sahlberg, 2007). Such an approach integrates inquiry with development: at its core is the thinking professional. An inversion of current thinking about research organization and funding is needed to emulate the success of the Finnish model; an integration of research and practice is required. Researchers, teacher educators, and funders need to work out how funding for education research can be shifted from large, putatively disseminative programs to small, teacher-sized, class-sized, and school-sized pockets of funding. Instead of large-scale funding for "What works?" evaluations, small "good practice scholarships" could be offered to practicing teachers working under the tutelage of university lecturers to guide practitioners in action research and case study. One multimillion dollar "What works?" evaluation would fund a thousand such scholarships. Tying such scholarships more firmly to the postgraduate education of teachers would contribute to the culture of continuing professional development that operates with such success in Finland.

For there to be "research *in* practice, not research on practice," as Friedman (2006, p. 132) puts it, structured conversations need to be possible both on university campuses and locally within consortia of schools for those in receipt of good practice scholarships. Therein would be nurtured the communities of practice and inquiry about which we speak so much, and therein would lie also a basis for fulfilling the conditions of collaborative inquiry in which people in organizations are galvanized to work as coresearchers (Argyris & Schön, 1996) rather than submitting themselves to the role of passive—and potentially resentful—recipients. It would provide a way of working Dewey's (1923, p. 46) "unworked mine" of knowledge from classroom teachers. Thus, schools could become the social learning systems of which Wenger (2000) and others speak.

Cultivating a New Education Science

In the ninety-odd years since Dewey (1923) proffered his advice about *The Sources of a Science of Education,* education has not defined, cultivated, or celebrated its own distinctness or integrity as a science.[9] Lacking such definition, our science has been vulnerable to buffeting over the years from scientistic discourses that misunderstand not only education but also science itself. It has been prey to a host of interests and influences that promise science but portend, in the contexts I have discussed, nescience.

We need a new kind of science: education science. No longer should we borrow the inquiry premises of medics, plant scientists, or pharmacologists or even the precepts of psychologists, sociologists, or anthropologists. No longer should we see education's status reduced to a field of study nested within other disciplines. For if we do, we are perpetually in danger of attempting to emulate other sciences and doing what the celebrated neurologist Oliver Sacks (1996) calls, in the case of his own discipline, the "reverse of science" (p. xvii).

Education science should, I have argued, be about "slow and gradual independent growth of theories" (Dewey, 1923, p. 18). Education is about teachers teaching and students learning—it is about practice—and the science must develop its roots at this practical level. Arguing against the quantification that was, even at the time he was writing, increasingly characterizing research in education, Dewey suggested that "educational practices are . . . the final *test of value* of the conclusion of all researches" (p. 33).

For the educator, research is part of the craft. The two are intimately bonded, and it is in this bond that our science is made. Without this connection, research becomes merely a technical process done by others for teachers. And is it any surprise that this is resisted and rejected by the practicing teacher? (It is resisted, incidentally, in the way that it is not by the medic, because medical science is the right fit for the questions being asked in medicine [Booth, Hargreaves, Bradley, & Southworth, 1995].[10]) The conjoining and interconnection of teaching and research is the essence of the dynamic between theory and practice. Without this interconnection, there is forever going to be the gulf between research and practitioner communities noted by Huberman (1999); there will forever be disillusion, indifference, and detachment.

In his book *The Craftsman*, Richard Sennett (2009) provides a useful insight from professional work elsewhere. He suggests that much of the lack of purpose, the disillusion, the anomie of modern life comes from the dismissal of thoughtful involvement from the practice of a job. Too many of today's jobs are rendered colorless and hollow, argues Sennett, by the introduction of procedures that remove the pride and dedication to improvement—separating the research from the practice, if you like—from professionals' craft. The consequence is in professional skills that are offered peremptorily and unimaginatively. Because of the separation of thoughtful involvement about improvement—research, in other words—from practice, the desire to perfect and improve has more or less evaporated from many domains of life. That conjoining of head, hand, and heart has all but gone.

The press for the disseminative, top-down, knowledge-based, folk-science model of educational inquiry has just this kind of effect. It distances inquiry from the work itself; it uncouples research, inquiry, and scholarship from the practice of the teacher. The effect of this disengagement is not only to restrict research to questions about effectiveness but also to constrain opportunities for participation by practitioners in education decision making: practitioners are levered into unwilling trysts with researchers wherein researchers, administrators, and policy makers set the direction.

The type of thinking that leads to this state of affairs is based, as much as anything, on what Sennett (2009) calls "the tidy mind"—which is surely part of the apocrypha of folk science. He notes that the mess of work—of practical endeavor mixed up with research and development—is anathema to the tidy mind. Practical work does "something distasteful to the tidy mind, which is to dwell temporarily in mess—wrong moves, false starts, dead ends" (p. 161). But the fact is that the teacher creates mess as a way of understanding work. In this, there is a practical process of conjectures and refutations, and this process should be celebrated, not denied. In it is the heart of our science.

Our science must take account of the reality of education endeavor, binding itself intimately with practice and enabling thinking and reflection on the lineaments and interstices of individual practice. Singular and shared understandings of such practice lie at the heart of our inquiry, and we should strive to forge a new science of education based on such understandings.

Notes

1. Perhaps the greatest natural scientist of all time, Einstein was an enthusiastic proselytizer of freedom and openness in discussion about science. He did much to demystify its processes, clear up misunderstanding about the ways that it worked, and expose hollowness in the claims of those who would make inappropriate claims about its methods and the prospects for its contribution.
2. Despite increases in computing power of many orders of magnitude since Rescher's analysis, there is little evident improvement in modeling. The forecasters' advice is now, "The weather beyond about a week ahead stretches even the most experienced weather forecaster. Complex numerical weather forecast models from the Met Office and the European Centre for Medium Range Weather Forecasting are run many times for the month (and season) ahead to build up a picture of the likelihood of different weather types affecting the UK" (Met Office/BBC, 2011). They may as well have added, "But none of it does any good."
3. Slavin is not alone here, of course. Reviews of the theory and practice of randomized experimental group designs are given by Thomas and Pring (2004) and Schoenfeld (2006).
4. The reason for the need for RCTs in pharmacology and medicine and not in, say, aeronautical engineering is that medical interventions exist in a strange epistemic netherworld where people often get better anyway. When jets crash, they rarely get better. How do you know that people have got better because of your drug? Answer: RCT. It's an excellent answer to a difficult problem in pharmacology.
5. Flyvbjerg (2006b) contends that a "focus on minutiae . . . directly opposes much conventional wisdom about the need to focus on 'important problems' and 'big questions'" (p. 377). Social science's answers to the latter, he suggests, are likely to be unhelpful for all of the reasons I have given in education's case. His reasoning leads him to suggest that "some of the best management schools . . . have understood the importance of cases over rules and emphasize case-based and practical teaching. Such management schools may be called Aristotelian; whereas schools stressing theory and rules may be called Platonic" (p. 372).
6. The local judgments are constituted from trial and error and as such are akin to Popper's (1971) *conjectures and refutations*. We can think of the conjectures and refutations establishing "schemata" after Piaget (1975), or "intuitions" after Simon (1983), or "dra-

matic rehearsals" after Dewey (2007) or "thinking tools" after Bourdieu (in Wacquant, 1989, 50), or "middle axioms" after Bacon (Wormald, 1993), or "heuristics" after Tversky and Kahneman (1974), or "constructs" after Kelly (see Bannister and Fransella, 1971), or "Generalities II" after Althusser (1979, 183–190), or "common interpretive acts" after Schatzman (1991, 304). These notions encapsulate much the same idea.

7. See Gadamer (1975, p. 269) for discussion about variations in our "horizons of meaning."
8. Johnston's (1985) very similar case study examination of reading failure finds reasons for this failure more in students' anxiety than in putative psychological deficits, where traditional educational and psychological science often seeks explanations.
9. This is partly because of uncertainty about whether education even has integrity or distinctness as a discipline. See Conant (1963), Seckinger (1964), and Loughran (2009).
10. Booth et al. (1995) note that while doctors see themselves in an environment of continuing learning and journal reading, teachers see themselves differently. Teachers, in the main, reject mainstream education research.

References

Althusser, L. (1979). *For Marx.* London: Verso.

Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method and practice.* Reading, MA: Addison Wesley.

Bannister, D., & Fransella, F. (1971). *Inquiring man: The psychology of personal constructs.* Beckenham, UK: Croom Helm.

Becker, H. S. (1998). *Tricks of the trade.* Chicago: University of Chicago Press.

Berger, P. L., & Luckmann, T. (1979). *The social construction of reality: A treatise in the sociology of knowledge.* London: Penguin.

Berlin, I. (1979). The divorce between the sciences and the humanities. In I. Berlin, *Against the current* (pp. 80–110). London: Hogarth Press.

Booth, M., Hargreaves, D. H., Bradley, H., & Southworth, G. (1995). Training of doctors in hospitals: A comparison with teacher education. *Journal of Education for Teaching, 21*(2), 145–161.

Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin, 54*(4), 297–312.

Canguilhem, G. (1994). The various models. In F. Delaporte (Ed.), *A vital rationalist: Selected writings from Georges Canguilhem* (pp. 41–47). New York: Zone Books.

Carlyle, T. (1829). Signs of the times. Retrieved from http://www.victorianweb.org/authors/carlyle/signs/signs1.html

Cochran-Smith, M., & Lytle, S. L. (2006). Troubling images of teaching in No Child Left Behind. *Harvard Educational Review, 76*(4), 668–697.

Conant, J. B. (1963). *The education of American teachers.* New York: McGraw-Hill.

Dewey, J. (1920). *Reconstruction in philosophy.* New York: Holt.

Dewey, J. (1923). *The sources of a science of education.* New York: Horace Liveright. Available at http://www.archive.org/details/sourcesofascienc009452mbp

Dewey, J. (2007) *Human nature and conduct: An introduction to social psychology.* New York: Cosimo.

Einstein, A. (1941). Science, philosophy and religion: A symposium. In *The Conference on Science, Philosophy and Religion in Their Relation to the Democratic Way of Life.* Retrieved from http://www.update.uu.se/~fbendz/library/ae_scire.htm

Einstein, A. (1936). Physics and reality. *Journal of the Franklin Institute, 221*(3), 349–382.

Fenstermacher, G. D. (1994). The knower and the known: The nature of knowledge in research on teaching. *Review of Research in Education, 20,* 3–56.

Fish, S. (1989). *Doing what comes naturally*. Oxford: Clarendon Press.

Flyvbjerg, B. (2001). *Making social science matter.* Cambridge: Cambridge University Press.

Flyvbjerg, B. (2006a). Five misunderstandings about case-study research. *Qualitative Inquiry, 12*(2), 219–245.

Flyvbjerg, B. (2006b) Making organization research matter: Power, values and phronesis. In S. R Clegg, C. Hardy, T. Lawrence, & W. R. Nord (Eds.), *The Sage handbook of organization studies* (pp. 370–387). Thousand Oaks, CA: Sage.

Friedman, V. J. (2006). Action science: Creating communities of inquiry in communities of practice. In P. Reason & H. Bradbury (Eds.), *Handbook of action research (concise paperback edition)* (pp. 131–143). Thousand Oaks, CA: Sage.

Gadamer, H. G. (1975). *Truth and method.* New York: Seabury Press.

Gardner, M. (1996). *The night is large: Collected essays, 1938–1995.* New York: St. Martin's Press.

Haack, S. (2003). *Defending science—within reason: Between scientism and cynicism.* Amherst, NY: Prometheus Books.

Hennessy, S., Warwick, P., & Mercer, N. (2011). A dialogic inquiry approach to working with teachers in developing classroom dialogue. *Teachers College Record, 113*(9), 1906–1959.

Hirsch, E. D. (1976). *The aims of interpretation.* Chicago: University of Chicago Press.

Holton, G. (1995). The controversy over the end of science. *Scientific American, 273*(4), 168.

Huberman, M. (1999). The mind is its own place: The influence of sustained interactivity with practitioners on educational researchers. *Harvard Educational Review, 69*(3), 289–319.

Johnston, P. H. (1985). Understanding reading disability: A case study approach. *Harvard Educational Review, 55*(2), 153–177.

Kaestle, C. F. (1993). The awful reputation of education research. *Educational Researcher, 22*(1), 26–31.

Lagemann, E. C. (2000). *An elusive science: The troubling history of education research.* Chicago: University of Chicago Press.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific fact.* Beverly Hills, CA: Sage.

Leakey, L. B., & van Lawick, H. (1963). Adventures in the search for man. *National Geographic, 123*(1), 132–152.

Levin, J. R., & O'Donnell, A. M. (1999). What to do about educational research's credibility gaps? *Issues in Education, 5*(2), 177–230.

Loughran, J. (2009). Is teaching a discipline? Implications for teaching and teacher education. *Teachers and Teaching, 15*(2), 189–203.

Louis, K. S. (1992). Comparative perspectives on dissemination and knowledge use policies: Supporting school improvement. *Science Communication, 13*(3), 287–304.

Lynch, M. (1985). *Art and artifact in laboratory science.* London: Routledge.

MacIntyre, A. (1985). *After virtue: A study in moral theory.* London: Duckworth.

Maclure, M. (2009, September). *The offence of theory.* Paper presented to the British Educational Research Association Annual Conference, Manchester.

Marton, F., & Booth, S. (1997). *Learning and awareness.* Mahwah, NJ: Lawrence Erlbaum.

Massumi, B. (2002). *Parables for the virtual: Movement, affect, sensation.* Durham, NC: Duke University Press.

Medawar, P. B. (1982). *Pluto's republic.* Oxford: Oxford University Press.

Met Office/BBC. (2011). UK forecast monthly outlook. Retrieved from http://www.bbc.co.uk/weather/2635167

Mosteller, F., & Boruch, R. (Eds.). (2002). *Evidence matters: Randomized trials in education research.* Washington, DC: Brookings Institute.

Nadel, S. F. (1957). *The theory of social structure.* London: Cohen & West.

Newell, A., & Shaw, J. C., (1957). Programming the logic theory machine. In *Proceedings of the 1957 Western Joint Computer Conference* (pp. 230–240). New York: Institute of Radio Engineers.
Norman, G. (2003). RCT = results confounded and trivial: The perils of grand educational experiments. *Medical Education, 37*(7), 582–584.
Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*(2), 137–148.
OECD/UNESCO-UIS. (2003). Literacy skills for the world of tomorrow. Retrieved from http://www.pisa.oecd.org/dataoecd/43/9/33690591.pdf
OECD. (2010). PISA 2009 results: Executive summary. Retrieved from http://www.pisa.oecd.org/dataoecd/54/12/46643496.pdf
Ojala, M. (2005). How Finland is researching early childhood education. In B. Spodek & O. N. Saracho (Eds.), *International perspectives on research in early childhood education* (pp. 79–117). Charlotte, NC: Information Age.
Piaget, J. (1975). *A child's conception of the world*. Totowa, NJ: Rowman & Littlefield.
Popper, K. (1971) Conjectural knowledge: My solution of the problem of induction. *Revue Internationale de Philosophie, No. 95–96, Hypothèse et Confirmation, 25*(1–2), 167–197.
Ravitch, D. (2000). *Left back: A century of failed school reforms*. New York: Simon & Schuster.
Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.
Reid, K., & Valle, J. (2004). The discursive practice of learning disability: Implications for instruction and parent-school relations. *Journal of Learning Disabilities, 37*(6), 466–481.
Rescher, N. (1998). *Predicting the future*. Albany, NY: SUNY Press.
Ryle, G. (1990). *The concept of mind*. London: Penguin.
Sacks, O. (1996). *An anthropologist on Mars*. London: Picador.
Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of Educational Policy, 22*(2), 147–171.
Schatzman, L. (1991). Dimensional analysis: Notes on an alternative approach to the grounding of theory in qualitative research. In D. R. Maines (Ed.), *Social organisation and social process: Essays in honor of Anselm Strauss* (pp. 303–314). New York: Aldine.
Schoenfeld, A. H. (2006). What doesn't work: the challenge and failure of the What Works Clearinghouse to conduct meaningful reviews of studies of mathematics curricula. *Educational Researcher, 35*(2), 13–21.
Schön, D. A. (1987, April). *Educating the reflective practitioner.* Presentation to the Annual Meeting of the American Educational Research Association, Washington, DC. Retrieved from http://resources.educ.queensu.ca/ar/schon87.htm
Seckinger, R. K. (1964). Conant on education as a discipline. *History of Education Quarterly, 4*(3), 193–197.
Sennett, R. (2009). *The craftsman*. London: Penguin.
Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academies Press.
Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education, 41*(4), 455–470.
Simon, H. (1983). *Reason in human affairs*. Oxford: Basil Blackwell.
Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher, 31*(7), 15–21.
Taleb, N. N. (2008). *The black swan: The impact of the highly improbable*. London: Penguin.
Thagard, P. (1998). Ulcers and bacteria I: Discovery and acceptance. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biology and Biomedical Sciences, 29*(1), 107–136.
Thomas, G. (1997). What's the use of theory? *Harvard Educational Review, 67*(1), 75–105.

Thomas, G. (2007). *Education and theory: Strangers in paradigms.* Maidenhead, UK: Open University Press.

Thomas, G. (2010). Doing case study: Abduction not induction; phronesis not theory. *Qualitative Inquiry, 16*(7), 575–582.

Thomas, G. (2011). The case: Generalization, theory and phronesis in case study. *Oxford Review of Education, 37*(1), 21–35.

Thomas, G. & Pring, R. (Eds.). (2004). *Evidence-based practice in education.* Maidenhead, UK: Open University Press.

Tooley, J., & Darby, D. (1998). *Educational research—a critique.* London: Office for Standards in Education.

Turney, J. (2004, April 23). Private eye with "helps to inquiry" [Review of the book *Defending science—within reason: Between scientism and cynicism,* by S. Haack]. *Times Higher Educational Supplement.* Retrieved from http://www.timeshighereducation.co.uk/story.asp?storyCode=188226&sectioncode=39

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases, *Science, 185*(4157), 1124–1131.

U.S. Department of Education. (2004). *Toolkit for teachers.* Washington, DC: Author.

Wacquant, L. D. (1989). Towards a reflexive sociology: A workshop with Pierre Bourdieu. *Sociological Theory, 7*(1), 26–63.

Weiss, C. (1995). The four "I's" of school reform: How interests, ideology, information, and institution affect teachers and principals. *Harvard Educational Review, 65*(4), 571–592.

Wenger, E. (2000). Communities of practice and social learning systems. *Organization, 7*(2), 225–246.

Wittgenstein, L. (1958). *The blue and brown books.* Oxford: Blackwell.

Wormald, B. H. G. (1993). *Francis Bacon: History, politics and science, 1561–1626.* Cambridge, UK: Cambridge University Press.

Wright, T. (2010). Learning to laugh: A portrait of risk and resilience in early childhood. *Harvard Educational Review, 80*(4), 444–463.

Wright Mills, C. (1970). *The sociological imagination.* London: Pelican.

Zeichner, K. (2007). Accumulating knowledge across self-studies in teacher education. *Journal of Teacher Education, 58*(1), 36–46.

Ziman, J. (1991). *Reliable knowledge.* Cambridge, UK: Canto.