The Impact of Educational Research on Mathematics Education

Dylan Wiliam
Department of Education and Professional Studies
King's College London
Franklin-Wilkins Building
150 Stamford Street
London SE1 9NN

To appear in: A. Bishop. M. A. Clements. C. Keitel. J. Kilpatrick & F. K. S. Leung (Eds.). *Second International Handbook of Mathematics Education*. Dordrecht, Netherlands: *Kluwer Academic Publishers* (2003).

Running head: Impact of research on mathematics education

Telephone: +44 (0) 20 7848 3153
Fax: +44 (0) 20 7848 3182
Email: dylan.wiliam@kcl.ac.uk

Abstract

The failure of educational research to impact teaching and learning has been lamented almost from the beginnings of educational research itself. In this chapter, I explore some of the reasons for this failure and attempt to suggest how educational research might impact the teaching and learning of mathematics more effectively in the future. I begin by arguing that classroom practices in mathematics education are, in fact, relatively stable and have changed little over recent decades. To try to understand why that is, I trace briefly the history of educational research and suggest that the major reason for the failure of research to impact practice is the emphasis on analytic rationality as the method to generate knowledge in education. By focusing on the nature of expertise in teaching, I argue that a focus on Aristotle's intellectual virtue of 'practical wisdom' would be more appropriate and suggest some ways in this change of focus could lead to a more productive relationship between research and practice.

Introduction

In an interview towards the end of his life, Neville Barnes Wallis (1887-1979), the aircraft designer and creator of the 'bouncing bombs' used in attacks on the Mohne, Eder, and Sorpe dams in the Ruhr valley during the Second World War, was asked about his formative experiences. He described his arrival at a new school (he would probably have been 11 or 13), where his class spent all their-mathematics lessons for the first two weeks trying to determine the most accurate value they could for the ratio between the circumference and diameter of a circle. Barnes Wallis spoke about how the teacher had awakened his curiosity to find things out; he felt sure that the quality of the teaching he received was a crucial factor in his later success.

In 1966, a national school leaving examination in mathematics in England gave students four hours to answer one question from five:

1 Discuss the relevance of matrices to networks. Illustrate by suitable examples.
2 Discuss "Relations" with special references to their representations. Illustrate by suitable examples.
3 Discuss the applications of sets to linear programming.
4 [After a definition and an example of a simple continued fraction] Investigate simple continued fractions.
5 Investigate either: Quadrilaterals: classification by symmetry,
or: Triangles and their associated circles.

These two examples raise important questions about the idea of 'progress' in mathematics education. Is mathematics teaching better now than it was when Barnes Wallis was at school? Are the tests and examinations that we set today better than those that were set in 1966?

To some, these questions are hardly worth asking; mathematics teaching must have improved over the last 100 years, and of course our assessments are better than those of 35 years ago. And yet to insist on the possibility of progress is to make a claim about the nature of mathematics teaching and learning. We do not, for example, ask whether musical composition has improved since Mozart, nor whether oil painting has improved since Rembrandt.

Whether or not we believe that mathematics education has improved over the last 100 years, there can be little doubt that is has changed. Describing the nature of those changes, however, and tracing the extent to which they can be attributed to the results of educational research, are extremely difficult.

Over the last twenty years, there have been several attempts to review and synthesize research in mathematics education. In 1988 and 1989, the National Council of Teachers of Mathematics (NCTM) published its 'Research Agenda for Mathematics Education' (Charles & Silver, 1989; Grouws, Cooney, & Jones, 1988; Hiebert & Behr, 1988; Sowder, 1989; Wagner & Kieren, 1989) in order to "direct research efforts towards important questions" and "to encourage the development of support mechanisms essential to collaborative chains of enquiry" (Grouws et al., 1988 p. v). Three years later, the NCTM published its *Handbook of Research on Mathematics Teaching and Learning* (Grouws, 1992). Two of the main aims of the handbook echoed those of the Research agenda: "to synthesize and reconceptualize past research, [and] suggest areas of research most useful to advancing the field" (p. ix). In addition, a third goal was "where appropriate, provide implications of research for classroom practice". Because of their origins, these resources inevitably focused primarily on research from the Anglophone world, and in particular from the USA. In response to this focus as well as to address the need for synthesis and reconceptualization, the first *International Handbook of Mathematics Education* (Bishop, Clements, Keitel, Kilpatrick, & Laborde,

1997) made explicit efforts to cross national boundaries and to bring in research not published in English. Most recently, the *Handbook of International Research in Mathematics Education* (English, 2002) was initiated "in response to a number of recent global catalysts that have had an impact on mathematics education and mathematics education research" (p. 3) and aimed to "be proactive rather than reactive in examining the emerging and anticipated problems in our field" (p. 4).

These collections have provided an extraordinary level of scholarly synthesis of the research that has been undertaken in mathematics education and have been important in identifying priorities for future research. We have made huge progress in our understanding of "how, where and why people learn or do not learn mathematics" (Begle & Gibb, 1980, p. 8), and yet this knowledge about 'what works' in mathematics education does not appear to have had any great impact on the way that mathematics is taught and learnt around the world.

Of course, the observation that research in education has failed to affect practice is nothing new (see Kilpatrick, 1992), but in this chapter, I want to explore some of the reasons for this failure and to suggest how educational research might influence the practices of learning and teaching mathematics. To do so, I first discuss briefly some of the large-scale changes in mathematics education over the last twenty years or so. My argument here is that mathematics education (at least as it is practiced in classrooms) has changed very little over a considerable period. In order to try to understand why that is so, I then review the history of educational research and suggest that the reason that educational research has impacted so little on practice (in mathematics and in other subjects) has been due to the focus on analytic rationality as the prime method of educational research. In the subsequent section, I suggest that a focus on 'practical wisdom' rather than universal truths is more appropriate for social sciences such as education, and in the concluding section, I suggest how this change in focus might be brought about.

Large-scale changes in mathematics education

Attempting to describe changes in mathematics education across any period of time is surely an unwise endeavour. The changes will be different in each nation, and even within each nation, states and provinces often have very different education systems. More importantly, educational provision is rarely homogenous, so that even within a state or province, it will be impossible to describe in any meaningful way what is provided under the broad heading of 'mathematics education'.

A second difficulty is that the changes in mathematics education that take place within any one system are not in any sense a steady progression towards an optimum. For example, consider the 'modern mathematics' movement that began with the Royaumont Seminar in 1959 (Organisation for European Economic Co-operation Office for Scientific and Technical Personnel, 1961). Although not completely without benefit, the assumption that the epistemological hierarchies developed by mathematicians would provide the most suitable didactical hierarchies for students between the ages of 5 and 19 now seems like one of the craziest ideas to take widespread hold in education. There are few today who would argue today that the 'new math' ideas should be tried again (see Moon, 1986, for a history of the controversy). But if we are to learn from history (so that, in the words of George Santayana, we are not condemned to repeat it), we need to understand whether, even with hindsight, the idea of modern mathematics was always wrong, or was it, again with hindsight, right for the 1960s, but not for the 21st century?

Clearly, there is not enough space in this handbook, let alone this chapter, to attempt to cover the whole of research in education and its impact on mathematics education. Therefore, I will instead discuss some of the 'revolutions' that have been claimed in mathematics education. Although inevitably a subjective view, I hope to show that the changes in what actually happens day-to-day in mathematics classrooms (as opposed to

what happens across educational systems) over the last twenty or so years have actually been quite modest.

The first revolution I wish to consider is the 'technological revolution', which is perhaps the most visible change in mathematics education, at least in classrooms in Europe and North America. There can be little doubt that the use of information and communications technology (ICT) in mathematics classrooms—specifically of hand-held calculators and computers—is increasing, but the total impact of ICT on the teaching and learning of mathematics is still very limited.

The Third International Mathematics and Science Study (TIMSS) found that in the fourth grade (students typically 9 or 10 years old), there was a huge variation in the use of calculators in mathematics classes (Mullis, Martin, Beaton, Gonzalez, Kelly, & Smith, 1996, p. 178). In several education systems, over 90% of students sampled said that they never used calculators in mathematics classes, while in Scotland, only 5% never used calculators. The median figure across the 26 systems was 79%. In grade 8, calculator usage was much greater, and yet in half the countries studied, the majority of students said that they used calculators "once in a while" or "never" (Beaton, Mullis, Martin, Gonzalez, Kelly, & Smith, 1996, p. 166). Establishing the quality of the use of calculators in schools is much more difficult than establishing its extent, but in concluding a survey chapter on this issue, Ruthven (1997) stated that "calculators are confined to the margins of classroom life: casually used, primarily instrumentally, and often uncritically" (p. 464).

Because of their greater cost, computers are inevitably less widespread than calculators, but many governments have made significant efforts to ensure access to computers in schools. Despite the huge investment in equipment, Balacheff and Kaput (1997) state that, in terms of 'market penetration' in the use of computers "the expectations of thirty years ago have surely not been met" (p. 470). TIMSS found that in the countries on which dependable data were available, the proportion of eighth-grade mathematics teachers saying that computers were never, or almost never, used in their classrooms to solve exercises or problems ranged from 38% to 99%, with a median value of 89% (Beaton et al., 1996, p. 167), and in grade 4, the median value was 93% (Mullis et al., 1996, p. 179). In half the countries studied, at least 85% of 8th-grade students and 80% of 4th-grade students said they never used computers in lessons at all. Only in 5 of 26 countries did at least 10% of grade 4 students say they used computers in "most lessons", whereas in only 5 out of 44 countries did 10% of grade 8 students say they used computers "pretty often" or "always" in mathematics lessons. In other words, in the vast majority of the mathematics classrooms surveyed in TIMSS, computers were almost never used to solve exercises or problems, and rarely used for anything else.

And, as with calculators, we should also be concerned with the quality of the use of computers, as well as just the extent. Although there is substantial evidence of the utility of computer-based learning environments for teaching mathematics (what Balacheff and Kaput, 1997, call 'epistemological penetration', p. 470), studies of the impact of the use of computers suggest that they are not well integrated into classroom practice (Watson, 1993). As the title of a recent book by Larry Cuban has it, computers in education really have been "oversold and underused" (Cuban, 2002).

The second major 'revolution' in mathematics education that I wish to consider is the 'constructivist revolution'. Before this revolution, the standard psychological models of learning were based on information processing (derived from the idea of the mind as a computer) and before that, in Gestalt psychology, on learning as re-organisation or re-structuring. In terms of teachers' current practice, however, one of the most influential models of thinking appears to be one of the earliest: associationism (Mayer, 1977). In this view, learning consisted of strengthening links between stimulus and response through reinforcement, and failure to learn indicated a failure adequately to form or reinforce the correct links between stimulus and response. From such a perspective

errors are random failures in the chain of associations, to be corrected by reinforcement of the correct links. In contrast, constructivist approaches to psychology hold that learning is not the passive acquisition of associations between stimuli and responses, but is rather the result of an active process of sense making on the part of the learner. From such a perspective, the errors that learners make are not random, but, to an extent at least, predictable (Booth, 1984; Hart, 1981, 1984; Kerslake, 1986).

In the mathematics education research community, it seems that 'we are all constructivists now'. And yet, if one observes practice in mathematics classrooms all over the world or looks at textbooks, the predominant activity appears to be the repetition of mathematical techniques through exercises. Even in the education of the youngest students, where 'child-centred' methods have taken greater hold, it seems as if teachers' predominant models of learning, at least for mathematics, are more associationist than constructivist. It is also worth noting that although the writing of Piaget was undoubtedly influential, the adoption of 'child-centred' practices in primary schools during the 1960s and 1970s was due at least as much to changing philosophical views about the nature of the child.

There have, of course, been success stories. The concern with social justice and the search for ways of ensuring that historically disadvantaged groups enjoy greater success in mathematics have resulted in substantial achievements. The gap between female and male achievement in mathematics has been substantially narrowed over the last 50 years (Feingold, 1988; Friedman, 1989; Hyde, Fennema, & Lamon, 1990; Linn, 1992), and in many countries, females now outperform males up to age 18. In this context, however, it is worth noting that socio-economic disadvantage continues to have a profound negative impact on mathematics achievement, and in many countries, the achievement of students from minority ethnic communities is also cause for concern.

Inevitably, the foregoing examples are at best merely suggestive. Nevertheless I hope that even those readers who believe there has been change will agree that such changes have been limited, fragile, and highly vulnerable to changes in government policy.

What are we to make of this lack of significant change? Can lack of change be construed as success? Well, in many cases, yes. In countries that are moving from a system of education for only the elite to a system of mass education, or extending universal education beyond the primary school years, then just maintaining the status quo is an impressive achievement. Similarly, maintaining teacher quality in a changing social order, where the status of teachers and their relative monetary rewards are diminishing, is also an achievement.

Given the tremendous advances in what we know about successful mathematics teaching and learning, however, the huge gulf between what are conceived of as 'good mathematics classrooms' by the research community and the experience of most learners is a severe indictment of the failure of educational research to have an impact on practice in mathematics education. In the next section, I will explore some of the reasons for this failure.

The (troubling) history of educational research

Although the amount of money spent on educational research in most countries has been only a tiny fraction of the total expenditure on education (typically between one-third of a percent and one percent in Europe and North America), the large size of education budgets in most countries has meant that a large amount of money has undoubtedly been spent on educational research. Although there have been some small successes (Tyack & Cuban, 1995), the impact of research on practice is very hard to discern (Travers, 1983). And, notes William Reese (1999), "Nothing like a comprehensive understanding of the overall problem seems likely to appear on the horizon anytime soon" (p. 11).

The failure of educational research to have any real impact on educational practice in general, and on mathematics education in particular, has been lamented for many years (see, e.g., Kilpatrick, 1992). In 1945, J. Cayce Morrison, assistant commissioner for research in the State Education Department in New York, said that there was "too wide a gap between research at its best and much of its practice in education" (Morrison, 1945, p. 243). By the 1970s, there was increasing acceptance that educational research had not had much of an impact either on the practice of teachers in schools (Clifford, 1973) or on policy making (Weiss & Vickers, 1992; Ravitch & Vinovskis, 1995).

Today, there are it seems two broad strands of criticism of educational research. The first is that educational research is unnecessary or irrelevant, perhaps best summed up by Donald Campbell, who suggested that one of the key questions in research is, "If you are wrong about this, who will notice?" Too often, in social science research, he suggests, the answer is, "Nobody" (Campbell, 1986, pp. 128–129). This phenomenon manifests itself either in a belief that expert practitioners already know 'what works' in mathematics classrooms and so novice teachers can learn all they need to know by watching experienced teachers, or that pedagogical practice will always be weak and that the solution lies in prescribing curricula and teaching methods in 'teacher-proof' schemes.

The second strand regards educational research as necessary, but of poor quality. Too often, it is said, educational research produces results that are ambiguous or contradictory (Johnson, 1987). This attitude is perhaps best summed up by Robert F Kennedy's furious reaction to the ambiguous evaluation of the impact of additional money provided for the education of socioeconomically disadvantaged students: "Do you mean that you spent a billion dollars and you don't know whether they can read or not?" (Lagemann, 2000, p. 202).

More recently, in Great Britain, the government became so frustrated by the 'irrelevance' of educational research that it commissioned a report from a group of social scientists from the Institute of Employment Studies at the University of Sussex. The report concluded:

> Our overall conclusion is that the actions and decisions of policy-makers and practitioners are insufficiently informed by research. Where research does address policy-relevant and practical issues it tends to:
> • be small scale and fails to generate findings that are reliable and generalisable;
> • be insufficiently based on existing knowledge and therefore capable of advancing understanding;
> • be presented in a form or medium which is largely inaccessible to a non-academic audience; and
> • lack interpretation for a policy-making or practitioner audience.
> This results at least in part from a research effort that is predominantly supply (i.e. researcher) driven. Furthermore, the research agenda tends to be backward rather than forward looking – following policy not prompting it. This is partly due to an emphasis on evaluation within much of the limited volume of government sponsored research, rather than exploration and development. It also reflects a dissonance between the policy making and research production cycles. (Hillage, Pearson, Anderson, & Tamkin, 1998, pp. xi–xii)

On those few occasions when research does produce unambiguous results, researchers are told that they are telling practitioners and policy makers what they already knew. If educational researchers could only agree how to go about research properly, we are told, educational research could join the elite club of 'hard' sciences producing reliable knowledge (these people have in the past rather unkindly been described as suffering from 'physics envy').

At the beginning of the last century, educational research, to the extent that it existed at all, was either historical or an aspect of philosophy. One of the earliest attempts to use empirical methods in educational research was the 'School Survey' movement in the USA. Beginning around 1910, this movement sought to gather 'objective evidence' about factors influencing the educational progress of school students. However, because

of the sheer diversity of the U.S. education system, with over 100,000 school districts each free to determine its own education policy, there was little agreement about the purpose or scope of education, and meaningful comparisons of educational outputs was almost impossible.

In *An Elusive Science* (whose subtitle is the title of this section), Ellen Condliffe Lagemann (2000) shows that the search for ways of producing high-quality research in education has been, in effect, a search for secure disciplinary foundations for the production of reliable knowledge. At first, philosophy and history provided those foundations, but around the turn of the century, these were supplanted by psychology, which dominates to the present day, although since the 1970s sociology and social anthropology have also been influential.

Lagemann argues that the failure of educational research to deliver what has been wanted has arisen from three main causes: the isolation and low status of educational research in the academy; its tendency to focus too narrowly on particular aspects of education rather than looking at education systems; and the weak governance and regulation of educational research. These three causes are of course intimately entwined.

In the USA, teaching had been regarded as 'women's work' since early in the nineteenth century, so that educational research was accorded low status by association. Lagemann also points out that being an applied subject served to marginalise education within the academic mainstream. No doubt partly in an attempt to raise its status, educational research attempted to emulate the hard sciences through the quantification of educational processes, which of course entailed focusing on those aspects of education that could be easily quantified. And whereas most teachers were female, most school supervisors and district administrators were male, so that the emerging field of educational research emphasised educational administration almost from the outset.

This lack of agreement not just about how to undertake educational research but also what should be researched continued to plague attempts to establish 'what works' in education over the next half-century. As Kilpatrick (1992) noted:

> Educational researchers were not satisfying the requests of practitioners that they provide useful information, they were not satisfying the funding agencies and their colleagues in the scientific disciplines that their work was valuable, and they were not satisfying their own expectations for what research should be. Far from living in a golden age, they seemed to be entering a depression. (p. 31)

However, Lagemann's history closes with an ironic twist. In the final quarter of the last century, educational research finally began to get on the right track with two key realisations. Firstly, the complexity of educational settings requires that insights from all of the 'foundation disciplines' (and not just one) be required to make progress in educational improvement. Secondly, it slowly became clear that centre-periphery models of dissemination are simply ineffective in education. The result was a blossoming of multi-disciplinary research involving teachers in real innovation and improvement. At the same time, however, it seems that the politicians gave up on educational research, and by 1991, federal funding for educational research in the USA stood at approximately one-third the level provided in 1971.

Lagemann's analysis is persuasive, and there are many who believe that the nature of research in the physical sciences can provide a useful model for creating thorough foundations for educational research. For example, in the USA, the Committee on Scientific Principles for Education Research has recently produced a report entitled *Scientific Research in Education* (Shavelson & Towne, 2002) identifying six principles underpinning the conduct of research in all scientific disciplines that, it is argued, are equally appropriate for educational research. Although the invocation of research in the physical sciences as an ideal towards which educational research should aspire may

have some merit, I want to suggest here that the goal of educational research as a science, in the same sense as physics is a science, is not just elusive, but impossible.

This claim is in part philosophical, but it is also in part an empirical claim. The phenomena that are studied in educational research (and indeed in all social sciences) are, in the first instance, far more complex than those that are studied by the 'hard' sciences—just imagine trying to set boundary conditions for the initial state of a typical classroom. In addition, however, it is important to realise the autonomy that individual students bring to lessons is not a problem with which physical sciences have to grapple. Bars of iron do not behave differently because someone has been testing them. Or more precisely, although bars of iron may behave differently depending on how they have been treated in the past (e.g., whether they have been annealed or subject to repeated stress and strain), we know what kinds of treatments matter, and we know how to find out in advance how the bar will perform under tests. Even those who believe that there is no such thing as free will and that all human behaviour at time T1 is actually determined by the state of the system at T0 have to concede that it is too difficult to specify the starting conditions precisely enough to determine the outcome. Chaos theory and, at a smaller scale, Heisenberg's uncertainty principle render Laplace's dream of being able to predict all behaviour from initial conditions a nonstarter.

So, if educational research has difficulty in producing (or as I have argued cannot produce) the 'reliable knowledge' (Ziman, 1978) produced by the physical sciences, what is to be done? In the next section, I argue that educational research needs to focus not on the kind of analytic rationality espoused (although not always practised!) in the physical sciences but instead on the pursuit of practical wisdom.

Rationality and reason in educational research

In his Nicomachean Ethics, Aristotle proposes that there are three primary types of intellectual virtue: *episteme*, *techne,* and *phronesis*. *Episteme* concerns knowledge of eternal universal truths and equates to the kinds of knowledge that is developed in the physical sciences, often expressed in the form 'know that' and 'know why'. This kind of knowledge is independent of context and persists today in English words like *epistemology* and *epistemic*. *Techne*, in contrast, is the study of how to bring things into a concrete form. It differs from *episteme* in that *episteme* is concerned with things that are the way they are of necessity (otherwise they would not be eternal truths), whereas *techne* deals with things that are contingent and variable. It is thus more about craft knowledge and 'know how' and persists in English words such as *technology* and *technique*. *Phronesis*, often translated as *prudence*, is the ability "to deliberate rightly about what is good and advantageous" (Aristotle, 1976, 1140). This idea, as Aristotle points out, is different from *episteme* since there is no point in debating things that are universally true—*phronesis* requires knowledge of particular circumstances. *Phronesis* is also different from *techne* since it is designed to move people to action rather than production. Aristotle's point here is that *techne* is product-oriented because the aim of the production is not the production itself but the product, whereas action is process-oriented—the end is doing *well*. This distinction inevitably brings in notions of ethics and values. Unlike *episteme,* which aspires to be value free, *phronesis* is based on practical value-rationality. As Flyvbjerg (2001) notes:

> By definition, phronetic researchers focus on values; for example by taking their point of departure in the classic value-rational questions: Where are we going? Is it desirable? What should be done?" (p. 130)

The choices we make are not right or wrong in an absolute sense, but good or bad according to some set of values and ethical principles. For a fuller discussion of the idea that ethical and value judgements are at the heart of educational research, see Lester and Wiliam (2002).

Since *episteme* deals with universal truths, it is independent of individual experience. Those with different experiences should be able to agree on the extent to which a particular claim is universally true. With *phronesis*, however, individual experience is crucial. Experience can result in an individual reaching universal truths, but that is very rare. More often, the result of experience is not a set of rules but an ability to make good decisions, without necessarily being aware of how these good decisions are arrived at. This ability might be called intuition, but too often, from a standpoint rooted in analytic rationality, intuition is rejected as being 'unscientific'. From the point of view of *phronesis,* intuition is an essential element of practical wisdom—the (albeit implicit) result of a large-scale synthesis of substantial experience.

Although *phronesis* is relevant only when there is no universal truth, that does not mean there are no general principles involved. *Phronesis* involves the practical wisdom to act well by the successful integration of general principles with detailed consideration of context and therefore must be based on substantial experience as well as the use of general principles. In this context, it is worth noting that the word *general* has two meanings in English. To a mathematician, something is generally true only if there are no exceptions (i.e., if it is universally true). In other contexts, however, something is generally true when it is true most of the time but when there are also known to be exceptions. The mathematicians' use of *general* corresponds to *episteme*, whereas the vernacular usage relates more to *phronesis*.

The distinction between *episteme* and *phronesis* is similar to the distinction between analytic and hermeneutic philosophy proposed by Richard Rorty (1979). Analytic philosophy pursues timeless truths, whereas hermeneutic philosophy is intended to help people deal effectively (well?) with issues here and now.

The physical sciences have succeeded because they have focused on *episteme*. But the social sciences have failed for precisely the same reason. For social science in general, and for mathematics education in particular, there cannot be universal truths, because successful action will always require the integration of general principles and specific contextual details. To choose a trivial example, if we want to find the perpendicular bisector of a line, we know that it can always be found by drawing a line between the intersections of circular arcs centred on the endpoints of the line. This statement is true for all lines in Euclidean geometry and is therefore an example of *episteme*. In contrast, how we teach students this technique is contingent, variable, and underdetermined, and therefore requires *phronesis*.

In his most recent book, *Return to Reason,* Stephen Toulmin (2001) makes a similar argument. He suggests that the attempt to emulate the physical sciences by focusing on analytic rationality as the only way for producing knowledge has been crippling to the social sciences. In particular, an emphasis on rationality has focused attention on those aspects of social phenomena that are most amenable to rational analysis. The result has been that the settling of issues that cannot be resolved by the application of universal rules is regarded as 'unscientific'. He suggests that a criterion of 'reasonableness'—research as a guide to action—would be more appropriate.

One of the examples that Toulmin quotes is the work of the U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which was asked to make recommendations about the use of young children in medical and behavioural research. During a six-month period, the Commission discussed the issue intensively and sought the views of others by holding public meetings. At the end of this time, they were able to frame a set of proposals and conditions for acceptable research practices on which they were in almost total agreement.

The Commissioners then went home to write statements about their reasons for approving the recommendations. Toulmin (2001) says,

> At this point Babel set in. Practically speaking, they were in agreement, and agreed what they agreed about. The thing they could not agree about was *why* they had agreed about it—what reasons they had for concurring in the recommendations. Their moral perceptions coincided, but the reasons they gave depended (at least partly) on their backgrounds. In short—and Aristotle would say the same—their shared certitude about their perceptions was greater than the certainty, or uncertainty, that each of them felt about the general systems to which they were professedly committed. (p. 132)

In other words, the Commissioners could agree about what would be 'good and advantageous' but could not agree about any universal rules underlying that agreement.

So not only does the search for universal rules underdetermine action as argued above, it can also prevent agreement about what is best. Had the Commissioners started from their foundational principles, they might never have secured agreement, and the process would certainly have taken longer. Perhaps the resulting consensus might have been even stronger, and longer lasting, but pragmatically the time available for reaching consensus is not unlimited. As I have argued above, I do not believe that universal rules will ever be of value in education, but even if such rules were possible in the future, they are beyond us now. Therefore, educational research ought to place more emphasis on the pursuit of practical wisdom to guide practitioners rather than on the search for absolute truths.

The superiority of intuitive rather than rule-based thinking is demonstrated by an experiment conducted on a group of paramedics (Klein & Klein, 1981) cited by Flyvbjerg (2001). Six short video extracts of a person administering cardiopulmonary resuscitation (CPR) were shown to three groups: experienced paramedics, students being trained as paramedics, and people who taught life-saving techniques. They were then asked which of the six they would choose to resuscitate them if they needed CPR. Five of the six video extracts were of inexperienced trainees just learning CPR, and the sixth was a highly experienced paramedic. Of the experienced paramedics, 90% chose the experienced paramedic, whereas only 50% of the students did so. Only 30% of the instructors, however, chose the experienced paramedic.

Flyvbjerg argues that we can understand this apparently paradoxical result by considering the five levels of expertise in learning proposed by Dreyfus and Dreyfus (1986). At the *novice* level, the individual is guided by rules that are applied irrespective of context. For example, the novice teacher tends to try to apply the same sets of rules to all the classes he or she teaches. The *advanced beginner* begins to take situationally specific factors into account, and personal experience is often relied on more than context-independent rules. As experience accumulates, however, the number of recognisable elements or 'chunks' increases and threatens to overwhelm the individual. For example, the need for the school teacher to attend to the learning needs of her or his students, controlling the behaviour of some while also trying to make sure he or she interacts equally often with female and male students, can lead to a feeling of 'plate spinning'—dashing from one imperative to the next to try to attend to all. The *competent performer* is characterised by performance in which conflicting priorities are resolved through the use of strategies usually derived from conscious problem-solving behaviour. In contrast, the *proficient performer* acts quickly and intuitively, often doing the 'right thing' without conscious awareness. In this context, it is important to realise that 'intuition' is used here not as some irrational prejudice but rather, as outlined above, as the result of the sedimentation and synthesis of vast amounts of experience. Finally, in the *expert*, the ability to act quickly and intuitively in a range of contexts and settings is unified into a 'feeling' of the right thing to do. The use here of an emotive term—'feeling'—is not coincidental. Experts 'feel' the best course of action, not just with their mind but also in their whole body. Expertise is therefore not the culmination of rationality but rather transcends it. Expertise involves going beyond what can be done through rationality. That is not irrational but meta-rational (i.e., beyond rationality).

Therefore, Flyvbjerg argues, it appears that the paramedic trainers identified the trained paramedics less successfully because they looked for paramedics who followed the rules that they themselves taught. In other words, they were looking for those at the level of competent performers rather than proficient performers or experts. If we accept that the classification proposed by Dreyfus and Dreyfus also applies to teaching, then it seems likely that the failure of educational research to have an impact on educational practice stems from a similar limitation.

If we accept that the prime (although not the only) purpose of educational research is the improvement of educational processes, then it can succeed only if research findings are taken up by teachers and incorporated into their practice. There are, of course, other ways that educational research can influence practice. Ideas about how children learn that were developed in the Concepts in Secondary Mathematics and Science (CSMS) project (Hart, 1981) fed into the construction of the first version of the national curriculum for mathematics in England and Wales (Department of Education and Science & Welsh Office, 1989) and these ideas continue to underpin the national curriculum even though it has been revised three times. Many of those involved in the preservice training of mathematics teachers are themselves actively involved in research, so research can feed into practice in that way. Research can also influence practice through the production of textbooks, although, as Clements (2002) notes, it rarely does so. If educational research is to have any lasting impact on practice, however, it must be taken up and used by practitioners.

The kinds of prescriptions given by educational research to practice have been in the form of generalised principles that may often, even usually, be right but in some circumstances are just plain wrong. Experts can see that a particular recipe is inappropriate in some circumstances, although because their response is intuitive, they may not be able to discern the reason why. What gets learnt by the practitioner is that the findings of educational research are not a valid guide to action.

But more often, research findings also run foul of the opposite problem: that of insufficient specificity. Many teachers complain that the findings from research produce only bland platitudes and are insufficiently contextualised to be used in guiding action in practice. Put simply, research findings underdetermine action. For example, the research on feedback suggests that task-involving feedback is to be preferred to ego-involving feedback (Kluger & DeNisi, 1996), but what the teacher needs to know is, "Can I say, "Well done" to this student, now?" Moving from the generalised principles produced by educational research to action in the classroom is not a simple process of translation.

In the past, this process has been called dissemination and is now more often called knowledge transfer—both interesting metaphors, suggesting that all that needs to be done is to "share the results" (English, Jones, Lesh, Tirosh, & Bussi, 2002, p. 805) with practitioners about the latest findings and they will be used.

If expertise transcends rationality, as I have argued above, however, then the process of knowledge transfer cannot be one of providing instructions to novices, advanced beginners, or competent performers in the hope that they will get better. Rather what is needed is an acknowledgement that what teachers do in 'taking on' research is not a more or less passive adoption of some good ideas from someone else but an active process of knowledge creation:

> Teachers will not take up attractive sounding ideas, albeit based on extensive research, if these are presented as general principles which leave entirely to them the task of translating them into everyday practice—their classroom lives are too busy and too fragile for this to be possible for all but an outstanding few. What they need is a variety of living examples of implementation, by teachers with whom they can identify and from whom they can both derive conviction and confidence that they can do better, and see concrete examples of what doing better means in practice. (Black & Wiliam, 1998, p. 15)

Conclusion

In this chapter, I have argued that while there may have been large-scale changes in education systems across the world (i.e., the extension of mass schooling), the changes in mathematics classrooms have been relatively modest (and not always for the better). Educational research has failed to affect practice because of an over-attention to analytical rationality, when a greater emphasis on 'practical wisdom' would have served us better as the primary goal of educational research. For too long, we have employed 'centre-periphery' models of dissemination in which researchers suppose that their role is first to work out the best way to teach mathematics and then (optionally) to tell teachers what to do. When teachers fail to 'listen' to researchers, researchers too often adopt habits similar to the Englishman abroad: talk louder, more slowly, and use simpler words. We have treated the 'translation' of research findings into practice as a simple, perhaps even technical, task, when in fact, research on the nature of expertise in teaching suggests quite the contrary (Berliner, Stein, Saberrs, Brown Claridge, Cushing, & Pinnegar, 1989).

Stokes (1997) has rightly pointed out that research should be inspired both by a quest for fundamental understanding and by consideration of its use (what he terms 'Pasteur's quadrant'). I am not suggesting that the quest for fundamental understanding is inappropriate in mathematics education research—we certainly don't yet know everything about 'what works' in mathematics teaching, But I do argue that we know a lot more about 'what works' than we know about how to support teachers in mathematics classrooms to improve their practice. Given the focus on analytic rationality as the criterion against which most educational research has been measured, perhaps that is not surprising. After all, researching the professional development of teachers appears to be much less tractable than researching the best way to learn algebra. But if mathematics education research is to affect practice, then 'considerations of use' need to be given a much higher priority.

If we *are* to take considerations of use more seriously, then we need to understand that teachers who manage to incorporate findings from educational research into their practice are not just 'translating' or 'transforming' knowledge. They are *creating* knowledge, albeit of a distinctive and specialised kind, by integrating the general principles that they can derive from research with the special, detailed knowledge they have of the contexts in which they work.

Fortunately, we have examples that suggest a way forward. Cognitively Guided Instruction (CGI) provided teachers with structures around which they could organise their knowledge of children's thinking, originally in addition and subtraction problems (Carpenter, Fennema, & Franke, 1996; Carpenter, Fennema, Peterson, & Carey, 1988; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989), which resulted in greater self-confidence and enhanced problem-solving ability. In the QUASAR project (Silver & Stein, 1996), 'resource partners' (typically researchers from local universities) assisted schools in improving their practice by helping them formulate their own approaches to curriculum development and professional development. In the Kent-Medway-Oxfordshire Formative Assessment Project (KMOFAP), teachers were supported in developing their own plans in developing their classroom assessment procedures to give them better information about students' learning and to involve students more in their own learning (Black, Harrison, Lee, Marshall, & Wiliam, 2002). What these three projects have in common is that they attend both to the process and the content of teacher development (Reeves, McCall, & MacGilchrist, 2001). They attend to the *process* of professional development through an acknowledgement that teachers need time, and support from colleagues, in order to reflect critically upon and to develop their practice. By themselves, however, these are not enough. Teachers also need ideas about the directions in which they can productively take their practice, and thus there is a need

for work on the professional development of mathematics teachers to pay specific attention to subject-specific dimensions of teacher learning (Wilson & Berne, 1999).

The problem with these three interventions is that they are highly resource intensive and could never, in their current forms, be geared up to provide impact on a national scale. What we therefore need are studies about the extent to which small-scale projects can be scaled up. Within the resources likely to be available, there is no possibility that such studies could establish 'what works' according to the criteria of analytic rationality, but they could just possibly do so if our criterion is one of reasonableness. Similarly, although Geraldine Clifford (1973) notes that proving that research affects teaching is "a near impossibility" (p. 3) within the framework of analytic rationality, if we adopt 'reasonableness' as our criterion, perhaps we can begin to make progress there too.

I am aware that my discussion of *phronesis* has excluded considerations of power, and this is a severe limitation. By founding our basis for action on value-rationality rather than on universal truths, we raise the immediate question of which people's values are to count? Although research cannot tell practitioners what to do, it can problematise all the taken-for-granted assumptions about our practices, such as what we do that serves to restrict opportunities to high-quality mathematics education to some groups of students (Moreno-Armella & Block, 2002; Tate & Rousseau, 2002). As Michel Foucault observed in respect of his work on prisons:

> It's true that certain people, such as those who work in the institutional setting of the prison ... are not likely to find advice on instruction in my books to tell them "what is to be done." But my project is precisely to bring it about that they "no longer know what to do," so that the acts, gestures, discourses that up until then had seemed to go without saying become problematic, difficult, dangerous. (quoted in Miller, 1993, p. 235)

A focus on *phronesis* can serve to open up debate so that other voices can be heard, thus enabling more conflicting stories to emerge. This process contrasts starkly with *episteme*, which seeks to close down debate by having the last word. In mathematics education research, there can be no objectivity:

> There is *only* a perspective seeing, *only* a perspective 'knowing;' and the *more* affects we allow to speak about one thing, the *more* eyes, different eyes, we can use to observe one thing, the more complete will our 'concept' of this thing, our 'objectivity' be". (Nietzsche, 1969, p. 119, emphasis in original)

In closing, I can do no better than echo the words of Bellah, Madsen, Sullivan, Swidler and Tipton (1985), by hoping that "the reader will test what we say against his or her own experience, will argue with us when what we say does not fit, and, best of all, will join the public discussion by offering interpretations superior to ours that can then receive further discussion" (p. 307).

References

Aristotle. (1976). *The Nicomachean ethics*. London, UK: Penguin.

Balacheff, N., & Kaput, J. J. (1997). Computer-based learning environments in mathematics. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 469–501). Dordrecht, Netherlands: Kluwer Academic Publishers.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's third international mathematics and science study*. Chestnut Hill, MA: Boston College.

Begle, E. G., & Gibb, E. G. (1980). Why do research? In R. J. Shumway (Ed.) *Research in mathematics education* (pp. 3–19). Washington, DC: National Council of Teachers of Mathematics.

Bellah, R. N., Madsen, R., Sullivan, W. M., Swidler, A., & Tipton, S. M. (1985). *Habits of the heart: Individualism and commitment in American life*. New York: Harper & Row.

Berliner, D. C., Stein, P., Saberrs, D., Brown Claridge, P., Cushing, K., & Pinnegar, S. (1989). Implications of research on pedagogical expertise and experience for mathematics teaching. In D. A. Grouws, T. J. Cooney, & D. Jones (Eds.), *Perspectives on research on effective mathematics teaching* (pp. 67–95). Reston, VA: National Council of Teachers of Mathematics.

Bishop, A. J., Clements, K., Keitel, C., Kilpatrick, J., & Laborde, C. (Eds.). (1997). *International handbook of mathematics education*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). *Working inside the black box: Assessment for learning in the classroom*. London: King's College London, Department of Education and Professional Studies.

Black, P. J., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: King's College London School of Education.

Booth, L. R. (1984). *Algebra: Children's strategies and errors*. Windsor, UK: NFER-Nelson.

Campbell, D. T. (1986). Science's social system of validity-enhancing collective belief change and the problems of the social sciences. In D. W. Fiske & R. A. Shweder (Eds.), *Metatheory in social science: Pluralisms and subjectivities* (pp. 108–148). Chicago: University of Chicago Press.

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively guided instruction: A knowledge base for reform in primary mathematics instruction. *Elementary School Journal, 97,* 3-20.

Carpenter, T. P., Fennema, E., Peterson, P. L., & Carey, D. (1988). Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education, 19*(5), 385–401.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*(4), 499-531.

Charles, R. I., & Silver, E. A. (Eds.). (1989). *The teaching and assessing of mathematical problem solving*. Reston, VA: National Council of Teachers of Mathematics.

Clements, D. H. (2002). Linking research and curriculum development. In L. D. English (Ed.) *Handbook of international research in mathematics education* (pp. 599–630). Mahwah, NJ: Erlbaum.

Clifford, G. J. (1973). A history of the impact of research on teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching* (pp. 1–46). Chicago: Rand McNally.

Cuban, L. (2002). *Oversold and underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.

Department of Education and Science & Welsh Office. (1989). *Mathematics in the National Curriculum*. London: Her Majesty's Stationery Office.

Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: Free Press.

English, L. D. (Ed.). (2002). *Handbook of international research in mathematics education*. Mahwah, NJ: Erlbaum.

English, L. D., Jones, G., Lesh, R., Tirosh, D., & Bussi, M. B. (2002). Future issues and directions in international mathematics education research. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 787–812). Mahwah, NJ: Erlbaum.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist, 43,* 95-103.

Flyvbjerg, B. (2001). *Making social science matter: Why social inquiry fails and how it can succeed again*. Cambridge: Cambridge University Press.

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematical tasks. *Review of Educational Research, 59*(2), 185-213.

Grouws, D. A. (Ed.). (1992). *Handbook of research on mathematics teaching and learning*. New York: Macmillan.

Grouws, D. A., Cooney, T. J., & Jones, D. (Eds.). (1988). *Perspectives on research on effective mathematics teaching*. Reston, VA: National Council of Teachers of Mathematics.

Hart, K. M. (Ed.). (1981). *Children's understanding of mathematics: 11-16*. London: John Murray.

Hart, K. M. (1984). *Ratio: Children's strategies and errors*. Windsor, UK: NFER-Nelson.

Hiebert, J., & Behr, M. (Eds.). (1988). *Number concepts and operations in the middle grades*. Reston, VA: National Council of Teachers of Mathematics.

Hillage, J., Pearson, R., Anderson, A., & Tamkin, P. (1998). *Excellence in research on schools*. London: Department for Education and Employment.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139-155.

Johnson, W. R. (1987). Empowering teachers: Holmes, Carnegie and the lessons of history. *History of Education Quarterly, 27*, 221-240.

Kerslake, D. (1986). *Fractions: Children's strategies and errors*. Windsor, UK: NFER-Nelson.

Kilpatrick, J. (1992). A history of research in mathematics education. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 3–38). New York: Macmillan.

Klein, H. A., & Klein, G. A. (1981). *Perceptive/cognitive analysis of proficient cardio-pulmonary resuscitation (CPR) performance.* Paper presented at the annual meeting of the Midwestern Psychological Association, Chicago.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254-284.

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

Lester Jr., F. K., & Wiliam, D. (2002). On the purpose of mathematics education research: Making productive contributions to policy and practice. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 489–506). Mahwah, NJ: Erlbaum.

Linn, M. C. (1992). Gender differences in educational achievement. In Educational Testing Service (Ed.), *Sex equity in educational opportunity, achievement, and testing: Proceedings of a 1991 ETS Invitational Conference* (pp. 11–50). Princeton, NJ: Educational Testing Service.

Mayer, R. E. (1977). *Thinking and problem-solving*. Glenview, IL: Scott-Foresman.

Miller, J. (1993). *The passion of Michel Foucault*. New York: Simon & Schuster.

Moon, B. (1986). *The 'new maths' curriculum controversy: An international story*. London: Falmer.

Moreno-Armella, L., & Block, D. (2002). Democratic access to powerful mathematics in a developing country. In L. D. English (Ed.) *Handbook of international research in mathematics education* (pp. 301–321). Mahwah, NJ: Erlbaum.

Morrison, J. C. (1945). The role of research in educational reconstruction. In N. B. Henry (Ed*.), American education in the postwar period: Part 2. Structural reorganization* (Forty-fourth Yearbook of the National Society for the Study of Education, pp. 238–265). Chicago: University of Chicago Press.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the primary school years: IEA's third international mathematics and science study*. Chestnut Hill, MA: Boston College.

Nietzsche, F. (1969). *On the genealogy of morals*. New York: Vintage Books.

Organisation for European Economic Co-operation, Office for Scientific and Technical Personnel (Ed.). (1961). *New thinking in school mathematics*. Paris: Organisation for European Economic Co-operation.

Ravitch, D., & Vinovskis, M. (Eds.). (1995). *Learning from the past: What history teaches us about school reform.* Baltimore, MD: Johns Hopkins University Press.

Reese, W. J. (1999). What history teaches about the impact of educational research on practice. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (pp. 1–19). Washington, DC: American Educational Research Association.

Reeves, J., McCall, J., & MacGilchrist, B. (2001). Change leadership: Planning, conceptualization and perception. In J. MacBeath & P. Mortimore (Eds.), *Improving school effectiveness* (pp. 122–137). Buckingham, UK: Open University Press.

Rorty, R. (1979). *Philosophy and the mirror of nature*. Princeton, NJ: Princeton University Press.

Ruthven, K. (1997). Calculators in the mathematics curriculum: The scope of personal computational technology. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 435–468). Dordrecht, Netherlands: Kluwer Academic Publishers.

Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The 'revolution of the possible' in mathematics instructional reform in urban middle schools. *Urban Education, 30*, 476-521.

Sowder, J. T. (Ed.). (1989). *Setting a research agenda*. Reston, VA: National Council of Teachers of Mathematics.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Tate, W., & Rousseau, C. (2002). Access and opportunity: The political and social context of mathematics education. In L. D. English (Ed.) *Handbook of international research in mathematics education* (pp. 271–299). Mahwah, NJ: Erlbaum.

Toulmin, S. (2001). *Return to reason*. Cambridge, MA: Harvard University Press.

Travers, R. M. W. (1983). *How research has changed American schools: A history from 1840 to the present*. Kalamazoo, MI: Mythos Press.

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.

Wagner, S., & Kieren, C. (Eds.). (1989). *Research issues in the learning and teaching of algebra*. Reston, VA: National Council of Teachers of Mathematics.

Watson, D. M. (Ed.). (1993). *The ImpacT report: An evaluation of the impact of information technology on children's achievements in primary and secondary schools*. London: King's College London, School of Education.

Weiss, C. H., & Vickers, M. (1992). Research impact on educational policy. In M. C. Alkin (Ed.), *Encyclopaedia of educational research* (Vol. 3, pp. 1093–1099). New York: Macmillan.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (pp. 173–209). Washington, DC: American Educational Research Association.

Ziman, J. (1978). *Reliable knowledge: An exploration of the grounds for belief in science*. Cambridge: Cambridge University Press.