

Manifesto for Evidence-Based Education

By Robert Coe, 1999

"Evidence-based" is the latest buzz-word in education. Before long, everything fashionable, desirable and Good will be "evidence-based". We will have Evidence-Based Policy and Evidence-Based Teaching, Evidence-Based Training - who knows, maybe even Evidence-Based Inspection.

But "evidence-based" is more than just trendy jargon. It refers to an approach which argues that policy and practice should be capable of being justified in terms of sound evidence about their likely effects. Education may not be an exact science, but it is too important to allow it to be determined by unfounded opinion, whether of politicians, teachers, researchers or anyone else.

The notion of "evidence" is not without problems. Many will say that one person's "evidence" may be another's opinionated nonsense. It is important not to deny these problems. However, in other areas, such as in the law or in medicine - even in science - the concept of evidence is potentially just as problematic, but has nevertheless been used successfully as a basis for decision making.

In the UK, and no doubt elsewhere, too many policies have been imposed on schools without adequate evidence about their likely effects and costs. It is arguably a waste of public money and professionals' time to impose policies without good evidence that they will lead to an improvement over what previously existed.

We need a culture in which evidence is valued over opinion, in which appropriate action (or inaction) is valued over just action for the sake of being seen to do something. By advocating such a culture, we hope to reclaim debates about policy and practice for the professionals who know most about them. In this way we hope to be able to do justice to the enormous responsibilities and hopes that are attached to education.

What do we mean by Evidence?

Evidence, like Motherhood and Apple Pie, is in danger of being all things to all people.

Everyone believes in the value of evidence and surely no one is against it, but what does it really mean? The following principles define more sharply a belief in Evidence-Based Education:

Sources Of Evidence

The only worthwhile kind of evidence about whether something works in a particular situation comes from trying it out. Arguments from theory are simply no match for something that has been tried and tested. Education has too many examples of plausible, convincing strategies for improvement that, when put into place, have either not had the desired effects or have had them, but at too high a price in terms of costs or unintended side effects. Schools and other educational

systems are just too complicated for us to be able to make reliable predictions in the absence of realistic experimental trials. In the language of research, that means doing well controlled field experiments.

Much of the research that is done in education takes the form of surveys or "correlational" research: looking for relationships within existing systems. This kind of research can provide valuable insights into how different features of a situation are related, but it cannot tell us what the effects will be of making changes. For example, correlational studies of class size show that children taught in large classes learn as much as, and often more than, children in smaller classes. A few moments' thought can provide explanations for this apparently anomalous result. When schools have a choice, do they not put more manageable children into larger classes, thereby allowing more difficult pupils to be taught in smaller groups? And are not more popular, "successful" schools likely to be oversubscribed and so have larger classes than those with lower levels of achievement?

However, if we were to take this correlational research as a basis for action, we might be in danger of advocating large classes as a way of raising achievement. This, of course, is exactly the error committed by much of the research on "school effectiveness". Correlations between factors such as "high expectations" and high achievement are taken to imply that encouraging teachers to expect more would lead to a rise in achievement.

It is only by doing an experiment in which children are randomly allocated to classes of different sizes that we could truly separate the effects of class size from all the other related factors, and so judge the likely effects of a policy of say, reducing class sizes from 35 to 20. When such experiments have been done, they have found that indeed the children in smaller classes do learn more - although the size of the difference is perhaps not as large as might have been hoped (Finn and Achilles, 1990; Blatchford and Mortimore, 1994). When experiments have been done to see whether achievement can be raised by raising expectations, it has been found that - at least for experienced teachers who know their material and their students - it is not really possible to raise their expectations artificially. So this plausible strategy turns out not to be very helpful in practice (Raudenbush, 1984).

The results of experiments can sometimes be disappointing, and this has led to some rejection of the experimental method of enquiry. However, this is a version of "shoot the messenger": if the method is sound but fails to demonstrate the success of a particular strategy, then perhaps the strategy is at fault. It is a strange kind of logic that would argue that if a strategy is unlikely to be successful it would be better not to know this before implementing it. Surely, it is always better to know. Educational researchers are often disappointed if they get negative results, but negative results, particularly if they were to prevent us from wasting time and money on ineffective policies, might actually be more useful than positive ones.

One further objection to the use of experiments is that they are hard to do because they require large samples in order to obtain a "significant" result. This, however, is a misconception. Using the technique of "meta-analysis" (Fitz-Gibbon, 1984, 1985) it is possible to combine the results of numerous studies, of whatever size, as if they were all parts of a single overall study. Meta-analysis can give estimates of how big an effect is, based on an "average" of the available

evidence. This "effect size" is an important concept in evidence-based education. Meta-analysis can also identify the factors on which the effect may depend: if some studies have found it and others not, meta-analysis can help to say why. This ability to combine results from separate experiments is a very powerful research tool. A small experiment in a single school could be quite hard to interpret; but if other schools have done the same experiment and the results are combined, the overall findings could be highly significant.

The notion of "evidence" is quite problematic. Some of the issues are acknowledged here:

Evidence is not value-free.

- Try to get a roomful of teachers, politicians or educational researchers to agree on the meaning of even something as apparently simple as "effective practice". One person's "effective practice" is another's "neo-liberal hegemony". A recent report commissioned by Ofsted on the state of educational research accuses much of it of being "partisan" - which indeed it is, although it seems a bit rich for anything to be called "partisan" by Ofsted. Opponents of the "evidence-based" philosophy cite the value-laden nature of all "evidence" as a major problem with the approach. But need this be a problem? All educational endeavour carries its own values and objectives. If we tried to make these explicit we might be able to have an honest and democratic discussion about what values are important, instead of passing off a concealed values debate as a debate about evidence. We may not agree about the particular values, but at least we will then know what we are talking about.

There are no universal solutions or quick fixes.

- Education is so complex that slight and subtle differences in context may make all the difference to the effects of some change in policy or practice. One of the things that makes teaching so interesting is that what works for me may not work for you, and even may not work for me the next time I try it. This makes it seem unlikely that the kinds of simple, universal strategies much beloved by governments will produce the improvements intended, although only by trying them can we really be sure. However, this complexity does not mean that there can never be any worthwhile evidence about anything. A more useful kind of evidence would be that which seeks to throw light on exactly which features of the context are important, to find the conditions on which the outcome depends. When we have this kind of evidence we will understand better which strategies are likely to be most effective in each situation.

Evidence is often incomplete or equivocal.

- One of the problems that commonly afflicts politicians is feeling the need to act, or at least to be seen to be acting, despite the absence of any clear evidence about what action is most appropriate. A more mature response in many areas of educational policy would be to acknowledge that we do not really know enough to support a clear decision. If the notion of using evidence as a basis for policy and practice is to take hold, then perhaps

we need to face up to this inadequacy and not make exaggerated claims for our limited knowledge. One of the implications of adopting evidence-based education would be that where evidence about some issue of importance is incomplete or ambiguous, we would not simply accept the fact, but seek to obtain the evidence we need. Equally where evidence is equivocal, we would need to explore the nature of the conflicting evidence and design further experiments to try to resolve it. Sadly, however, this almost never happens in educational research.

Evidence can be quite complex.

- In particular when statistical analyses are involved, evidence can be hard to interpret simply. Of course, it would be wrong to over-simplify, but it would be just as wrong not to make the evidence as accessible as possible. As good teachers know, most people can understand most things if they are explained well.

How Could Education Be "Evidence-Based"?

There are three main ways in which education could become more "evidence-based". These concern, firstly, the development of evidence-based policies, secondly, of evidence-based practice and, thirdly, the general promotion of a "culture of evidence".

Evidence-Based Policies

The need for evidence-based policies is clear. Politicians, government agencies and others repeatedly impose policy on schools with no - or at best limited - evidence about the likely costs and benefits. Change of any sort is always disruptive and should surely have to be justified. It would generally not be difficult or expensive to evaluate policies before they are imposed, and this should be routinely required.

There is an important difference between the kind of evidence that may justify advocating something as "good practice" and requiring it as policy. In the former case, teachers can make a judgement about whether something is applicable to their context or can adapt it to be most effective. Evidence derived from schools who have volunteered to try a particular innovation may well be relevant here: if it has worked for them, then it may well work for you. In the case of policy, however, where all schools are compelled to change what they are doing, the evidence to support it must be much stronger. It is not enough to know that something has worked in a selected group of volunteer schools. It may well be that schools who volunteer to take part in an initiative would have improved no matter what they did: identifying that there is a problem and being prepared to address it are often the largest parts of solving it. To justify imposing something as policy we must have evidence that it will work (or at least do no harm) in the hardest pressed, least amenable, most unlikely schools, who may well not volunteer for anything, as well as in schools that are already excellent and may reasonably feel no need to change.

Evidence-Based Practice

There already exists a sizeable body of evidence on "good practice", though there may be questions about how sound some of it is, how familiar it is to practitioners and the extent to which it actually influences practice.

I have already argued that for practice to be soundly based on evidence, that evidence must come from experiments in real contexts. "Evidence" from surveys or correlational research is not a basis for action. This is not to say that these kinds of research are worthless; far from it. In order to know what strategies are worth evaluating experimentally we need to have the kind of prima facie case for their effects that is most easily gained from correlational and survey research. However, we should not make the mistake of thinking it is all over when we have done this research. The only really sound evidence about what works comes from actually trying it.

There has been a good deal of controversial debate recently about the extent to which practitioners are aware of research evidence, and whose fault this is. "Dissemination" is now a major part of many research contracts, which means that researchers are expected to disseminate their findings to practitioners. Certainly, much educational research is not presented in a form (or a forum) that makes it very accessible to teachers.

But there is more to the issue of dissemination than this. One of the descriptions that has been applied to CEM monitoring systems is "distributed research" (Fitz-Gibbon, 1996). The intention has always been that while CEM may send performance information into schools, it is only those who work in the schools who can interpret it. This makes the whole activity into a research collaboration, since research is not just about measuring things, but, crucially, about interpreting the results.

In the same way, practitioners must be more than just the recipients of researchers' findings. Teachers' local knowledge makes them vital participants in the processes of framing research questions and in collecting and interpreting data. "Dissemination" must take the form of communication among practitioners as well as between practitioners and researchers - in both directions.

A Culture Of Evidence

The need for both policy and practice to be founded on evidence is part of a wider inadequacy in our approach to decision making across a range of social activity. It is part of a pre-scientific approach to social "science" that would look extremely incongruous in applied sciences such as engineering or medicine. No one would advocate building a particular structure, investing in an industrial process or using a medical treatment without good evidence to justify it. But it is part of the culture of social science that opinion can often have a status equal to - or even greater than - that of evidence as a basis for action.

We need to change that culture so that the question, "Where is the evidence?" becomes the first thing we think of when presented with a suggested change of practice or policy. The ease with which politicians, policy makers - and even teachers - have been able to get away with implementing their prejudices without even token consideration of the evidence, let alone engaging in a serious and informed debate about its quality and importance is a disgrace. Accountability must work two ways. Politicians want schools to be accountable for their performance, but they in turn must be accountable for the decisions they make - in the sense that they must be able to justify and account for them in terms of sound evidence.

References

- Blatchford, P. and Mortimore, P. (1994)
"The Issue of Class Size for Young Children in Schools: what can we learn from research?"
Oxford Review of Education, 20, 4, 411-428.
- Finn, J.D. and Achilles, C.M. (1990)
"Answers and Questions about Class Size: a Statewide Experiment."
American Educational Research Journal, 27, 3, 557-577.
- Fitz-Gibbon, C.T. (1984)
"Meta-analysis: an explication."
British Educational Research Journal, 10, 2, 135-144.
- Fitz-Gibbon, C.T. (1985)
"The Implications of Meta-analysis for Educational Research."
British Educational Research Journal, 11, 1, 45-49.
- Fitz-Gibbon, C.T. (1996)
Monitoring Education: Quality, Indicators and Effectiveness.
London: Cassell.
- Raudenbush, S. (1984)
"Magnitude of Teacher Expectancy Effects of Pupil IQ as a Function of Credibility of Expectation Induction: A Synthesis of Findings from 18 Experiments."
Journal of Educational Psychology, 76, 1, 85-97.