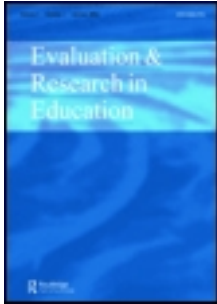


This article was downloaded by: [Universite Laval]

On: 16 January 2012, At: 13:15

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number:  
1072954 Registered office: Mortimer House, 37-41 Mortimer Street,  
London W1T 3JH, UK



## Evaluation & Research in Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/revr20>

## Evidence and Indicators: Dialogue, Improvement and Researching for Others

Hilary Constable & Robert Coe

Available online: 29 Mar 2010

To cite this article: Hilary Constable & Robert Coe (2000): Evidence and Indicators: Dialogue, Improvement and Researching for Others, Evaluation & Research in Education, 14:3-4, 115-123

To link to this article: <http://dx.doi.org/10.1080/09500790008666967>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or

damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Evidence and Indicators: Dialogue, Improvement and Researching for Others

*Hilary Constable*

*Faculty of Health, Social Work and Education, University of Northumbria at Newcastle, UK*

*Robert Coe*

*School of Education, University of Durham, Durham, UK*

The papers presented in this special double issue of *Evaluation and Research in Education* are taken from the first two conferences in what it is hoped will be a long series, entitled 'Evidence-based Policies and Indicator Systems' and held at the University of Durham in 1997 and 1999.

Although clearly not all of the papers from both conferences could be presented in this collection, we believe the variety and quality of those included are true to the flavour of the conferences. The rationale for the conferences themselves was eclectic. There was no attempt to impose coherence, merely a desire to invite interesting and respected scholars from a range of different fields and perspectives. Even the title of the conferences, 'Evidence-based Policies and Indicator Systems', seems to marry two unlikely bedfellows. In drafting the programme, intellectual stimulation counted for more than academic specialism. This, of course, makes the job of the editors of a collection such as this particularly difficult, but also particularly rewarding. We hope the reader will agree that a certain emergent coherence has been found and that the freshness and vitality of the ideas contained within this special edition are evident.

At the time when the first of these conferences was being planned, the notion of 'evidence-based policies' was yet to achieve the high profile that it has gained since then. 'Evidence-based' medicine was already relatively well established, with the formation of the Cochrane Collaboration in 1993 and, in the UK, the establishment of the Centre for Evidence-based Medicine in Oxford in 1995. However, debates about evidence-based policies, outside the field of healthcare, were still somewhat marginal.

By the time of the second conference, the social science research culture had arguably started to change. In particular, many of the funders of educational research, at least in the UK, had begun to use the expression 'evidence-based' in their calls for proposals. The presence at the second conference of so many researchers from around the world with interests in evidence-based policy provided the opportunity for an initial meeting of the Campbell Collaboration, a younger sibling of Cochrane, held in London immediately after the conference. The Collaboration was formally established at a meeting in Philadelphia in February 2000, with the aim of 'Preparing, maintaining and promoting the accessibility of systematic reviews of the effects of social and educational policies and practices' (<http://campbell.gse.upenn.edu>). Further developments in this area, as in many other areas of 'evidence-based policies and indicator systems' will no

doubt be of interest to participants in the third conference in the series, to be held in Durham on 4–7 July 2001. Readers will find details of this conference at <http://cem.dur.ac.uk/ebeuk>.

Despite the increasingly high profile of the concept of ‘evidence-based policy’ however, it remains unclear – or at any rate controversial – precisely what it means. Within the field of healthcare, evidence-based medicine has been defined as,

the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research. (Sackett *et al.*, 1996)

The ‘gold standard’ of ‘best evidence’, at least when applied to judgements about the effects of therapies, is taken to be ‘the systematic review of several randomised trials’. In areas such as education, social services and criminal justice, however, the tradition of experimentation, and particularly randomised experiments, is much less well established, and the ‘medical model’ remains controversial. It is certainly among the aims of the ‘evidence-based policies and indicator systems’ conferences to promote the best use of, and debate about, such methods. A number of the papers included in this volume address the issues of the appropriateness of different methodologies in providing ‘evidence’, and the means by which that evidence can influence policy and practice.

That there is any relationship between ‘evidence-based policies’ and ‘indicator systems’ is not perhaps obvious at first sight. The conferences in whose title they have been linked were hosted by Durham University’s Curriculum, Evaluation and Management (CEM) Centre, under the direction of Professor Carol Fitz-Gibbon. The main activity of this Centre is working with schools to provide comparative feedback on a range of performance indicators, notably on the ‘value added’ progress of their students. This work began in 1982 and currently over 7000 schools in the UK participate in one or more of the monitoring projects. The focus for these schools is on self-evaluation, based on fair comparisons with other schools derived from high-quality data. What distinguishes these schools, therefore, is not so much that they are interested in self-evaluation, but that they are particularly concerned with the quality of the evidence on which their judgements are based. Hence, it might not be stretching a point too far to describe their approach as one of ‘evidence-based’ management. Moreover, the availability of high quality measures of important outcomes provides those schools with an opportunity for relatively straightforward experimentation. Fitz-Gibbon (1996) has argued cogently for the need for both monitoring and experiments to guide practice and policy.

### Looking Outward and Learning from our Friends

This edition contains a range of papers from researchers working to improve evidence-based research and practice. They have been written in the context of a wider political drive to effect improvements in schooling. The authors are collectively working on ways in which indicator systems can be used to inform policy and the papers together form a balanced dialogue between researchers.

One important theme is the desirability of looking outwards and the conference theme was deliberately multi-disciplinary. Most of the papers presented here have an educational focus, but look in part to other fields of research like medicine, educational testing, data services and information science, and in part at the context in which evidence-based systems operate: the social and political systems. They also draw on experience from a range of countries and education systems in terms not only of practice and policy, but also the role of research in improving schooling. In this special edition of *ERIE* a wide range of perspectives has been drawn on, and this significantly enriches the picture.

Support and friendship are also contributions to research although they do not often get a mention in academic journals. The lasting value of a live interaction was plain to see at the two meetings from which the majority of these papers derive. Friendship, or learning from our friends, is another important theme in these papers. The longer and more heavily funded tradition of research in science, especially in medicine, means a vast reservoir of completed research. Papers in this edition show how contemporary technology can transform the body of existing research from what may feel like a deadweight into a living archive available to inform new research and action. The same papers remind us of a further group of friends, the information technologists. It is not so much that the arcane secrets of librarians are now more easily available to us through the use of information technologies, rather that those who employ these technologies have manifestly special skills from which we can – and need to! – learn.

### **Indicator Systems are Powerful**

A main focus of the papers is the development of research. It is hard to overestimate the potential of the research methods reported in these papers. At their best, the methods can identify key features of educational systems and substantiate their claims to be key features. They can identify and report on combinations of features. They can enable us to make sophisticated comparisons; to answer the challenge to compare like with like; to trace developments over time and to feel secure in the focusing of effort to improve schooling. They can help us identify key features in schooling and, what is more, to be confident in that identification.

The potential of indicator systems is powerful. The effects in terms of decisions and policies are not confined to the research community or even to professional educators but have effects in which the whole population is involved at some level or another. People want to know that evidence and indicator systems, which are used to inform policy, are good ones. This necessarily puts work on indicator systems into a position where not only the results, but also the activity itself attract interest and criticism.

The papers mostly consider what we can find out, but understanding the limits is important, not least in setting expectations. No one knows better than the practitioners the limitations of the various modes of evidence-based research. This edition shows researchers working together to improve research, exposing and discussing areas of uncertainty and controversy and identifying places for advance. One way and another all the papers in this edition both celebrate the achievements of research and at the same time point out the limitations.

Researchers and others need to know what this kind of research can do now,

what it may be able to do and the limits of what will be possible in the future. The reason all this matters so much is, of course, because it pertains to schooling where value positions cannot be avoided. There are two themes: one is perfecting the techniques and the other is keeping an eye on the relationship of the research with the whole picture of school improvement.

## The Ivory Tower

There are two major aspects of the relationship of research to its context which need to be considered. One is the relationship of research to change in practice, and the other is the relationship of research to the society of which it is a part. Unsurprisingly, these two aspects are themselves intertwined.

First the relationship of research to society. Who is research for and how do researchers view these audiences? The first question is easy; it is the second one that now impacts hard on the practice of research. The purpose of evidence and outcome-based research is, directly or indirectly, to improve schooling; the question of audience and stakeholders is neither so clear nor so simple.

Scientists have had some hard lessons to learn in the relationship of research to the society which both funds them and uses the results of their research. Researchers in education have the chance to learn from that experience: in areas like health (BSE and *E. coli* in industrial-scale meat production), genetics (GMOs in crops and medical intervention) and the environment (oil extraction and nuclear energy). The people who were initially excluded from these debates found ways of gate-crashing them. The non-experts have muscled their way into the debating hall in spectacular fashion. They have views about what should be researched, how and in what time scale, to say nothing of views about the use to which results are put. There is a paramount need to attend to the audiences of research. It is no longer possible to sit in a tower – even one constructed prosaically of pre-cast concrete blocks – and hope to perfect research techniques before launching them upon the world. The world wants to know.

In educational research the audiences are not limited to other researchers or to policy makers nor even to teachers and other professionals but include parents, children, in fact most of the population. We live in a society in which the education system is accountable to the citizenry, all the stakeholders and potential audiences. In contemporary society the way research relates to both policy makers and 'the people' is a necessary and urgent topic of debate for researchers. Research does not exist in isolation, neither is it static. As research advances so too do notions of accountability, democracy and participation.

Researchers cannot expect to be in a position to develop new methods away from the public gaze. They are active participants in what is potentially a very public debate. Right from the start people want to know what is going on, to be assured that the right things are getting attention, that the right methods are being used and the measurements pertain to things that are worthwhile. There are too many values embedded in notions of school improvement, and in research itself to leave citizens outside the debate. They must be inside. It is certainly a tall order when researchers have as their prime focus a comparatively technical matter. But it is worth reflecting that scientists have learned that if people living in a moderately democratic society such as the UK are not specifi-

cally engaged, then they will find ways of involving themselves in – perhaps even controlling – the work. A developing democracy is a tough and challenging habitat for practising researchers.

## Communicating with Stakeholders

There is an important duty on researchers to communicate and explain. This is especially important when it comes to explaining complex ideas and statistical techniques. Once again the previous experience of scientists in the work of the Committee on the Public Understanding of Science (COPUS) has a lot to offer researchers working in education.

Teachers occupy a special place in the debate. They are a special audience. It is their work that is being scrutinised and it is they who will, voluntarily or otherwise, make changes in their *modus operandi*. Teachers need to get beyond a position where they are dependent on authority-based recommendations or, put another way, the target for pot shots with researchers' recommendations. They need to enter the debate as informed individuals. What are the duties of researchers to educate this special audience in critiquing? How can or should teachers be part of the research? There is a need for teachers not only to be engaged by the aims and findings of the research, but also to be drawn into the research story as it unfolds. How can the research be made to work for teachers?

## Changing Schools

Another major aspect of evidence-based research is its relationship to change, knowing how to use the findings, how to implement change. To know how things are is only the starting point from which one can begin to understand how they should be changed or how to accomplish that change.

One critical pressure on research is the need for excellence at every level. But there are others including the need for timeliness (not always synonymous with 'speed') and economy (not always synonymous with 'thrift'). Children are being educated in indifferent or even bad schools at this minute and effective action must be both quick and economically sustainable. As with the war in Kosovo, so with low achievement by children from poor families: if enough people say something must be done it is in the interests of politicians to ensure it is done, but these acts may have unintended – and undesirable – consequences.

## Limits of Indicator Systems

Researchers working with indicator systems are fully aware that the indicators in current use are not perfect. They are keen to improve matters. In an ideal world educational systems would be released from the burden of excessive over-monitoring. In the UK for instance there are already school inspections by a government agency (OFSTED), publication of pupil achievement across the curriculum at ages 7, 11, 13 and 16 and appraisal of teachers, to say nothing of the monitoring systems discussed in this edition. On the face of it this is over-kill. However it is also a recognition that at present not one of the systems standing alone is up to the job.

Desirable or marketable knowledge in the future may have little to do with subject-based knowledge but may be about the ability to search systems or to

develop into a particular kind of citizen. It may be stretching the word 'curriculum' to try to fit the new needs to it. The jury is still out on whether or not increasingly sophisticated indicator systems will be able to handle this. On the more optimistic side, it is certainly true that there is much more that indicator systems can do in the way of looking at broader aspects of schooling. It is probable that indicator systems can spectacularly help develop the efficiency of current schooling systems. But there is a danger that the current measurements may slow down our ability to develop new curricula and may perpetuate certain sorts of knowledge and achievement. It is not clear that indicator systems will take education where it needs to go. What we do know is that what you measure is what you get.

## The Research Context

The first five articles concern the interface between researchers and their stakeholders: not only researchers and policy makers, but researchers and the citizenry.

The papers start with Brighton's 'Making Our Measurements Count'. His article is thought-provoking in the context of the problematic relationship between researchers and audiences. Brighton argues that debates about educational standards are too important to be left to teachers, academics and politicians. He goes on to take a broadly based critical stance on contemporary academic research techniques. He takes a swipe at a few of the habits of evidence-based researchers but worries unnecessarily that he might be seen as a Luddite simply because he makes an argument against multiple regression. Although for some people this may seem like arguing against motherhood and apple pie, it is clear that he is addressing much wider concerns and readers will see the more important point he is making. Brighton's case is that there is more than one way to do research but that there is no substitute for clear thinking.

By way of contrast Schweinhart shows what can be done when research and policy are in conversation in 'The High/Scope Perry Preschool Study: A Case Study in Random Assignment'. This paper is uplifting: it is optimistic about the case for intervention and it is positive about researching and creating generalisations. Schweinhart presents his view of the mounting evidence from research that the High Scope project works. Not a miracle but 'measured partial escapes from poverty and hopelessness' he says.

Schweinhart outlines the experimental basis of the work, in particular the random assignment of individuals to groups, and reminds us that this was possible because at the time there was a genuine doubt about whether or not the programme would have positive effects. Schweinhart exposes a number of issues with clarity. One of these is the difficulty of generalising to other programmes. Can one generalise from a half-day programme for pre-kindergarten/nursery children with weekly home visits to a whole-day programme with strong parent involvement? Is it possible to move from a teacher-pupil ratio of 1:5-6 to 1:8-10? In essence these are the decisions policy makers are faced with and they need all the help they can get from researchers. The policy makers need to know what they must do and the resources they require. To overspend is inefficient; but as Schweinhart says 'spending less per child than is needed to provide high-quality early childhood education, is not



only irresponsible, but also dishonest in that it offers false hope to people in need'.

Another point to come out of this paper is the need for short-term outcomes that are known to predict long-term outcomes. We cannot wait years to inform policy. In the third paper Wainer takes up the theme of conveying information to the audiences of research. Like Brighton he is concerned to keep a close eye on the relevance of questions. He is especially interested in the use of graphical presentation and takes the reader through a number of instructive examples. He includes examples of the relationship between expert thought and children's thought in Tectonic plates; between managers' thoughts and engineers' thought in the Challenger disaster and also of graphical deception in reporting the relationship between expenditure and (US) SAT scores and finally goes on to Simpson's paradox.

Wainer presents his argument with considerable elegance, drawing on historical examples. One is intrigued by the allusion to what happened when Playfair met Jefferson in France; one cannot fail to be moved by his account of 'the tragic convergence of fire and men in Mann Gulch'. His intellectual style, in which a dialogue is established between one branch of thinking and others, is an important part of the paper and one from which career academics might usefully learn.

Ridgway *et al.* in 'Problematising Evidence Based Policy and Practice' take the apparently straightforward example of evaluating the impact of a new mathematics curriculum and probe the ramifications. On the face of it and in their own words 'What could be simpler?' The authors start by discussing finding or devising tests which could do the job: both assessing different kinds of attainment and doing so in a way that could be compared fairly. They pursue this discussion through a list of further challenges to evidence-based policy and practice and conclude that our understanding and conceptualisation of larger systems is severely limited. Bringing the academic community closer to the core concerns of political decision making needs a new field of research focusing on macro-systemic change.

The paper from Hedges on 'Using Converging Evidence in Policy Formation: The Case of Class Size Research' addresses a fundamental issue for policy makers: the problem of different sources of evidence giving apparently different results. Hedges shows that, at least in the case of class size research, different methods give remarkably similar answers to the question 'What is the effect of reducing class size from 24 to 15?' Given that the different methods also have quite different strengths and weaknesses, the convergence of their results is a source of considerable confidence and credibility for the belief that reducing class size will lead to a modest increase in achievement. Hedges' finding is important, both for its methodological innovation and because of the level of political interest in this issue.

## Using the Evidence from Research

The next three papers make a set in which literature reviews are explored. They illustrate two important developments: the accumulation of a substantial body of existing research and the development of sophisticated techniques for handling information. There is now a substantial body of existing educational

research to look at, to learn from, evaluate and synthesise. We also want to make complex comparisons and we are in a hurry. We are at a moment in research well served by advances in information and communication technology (ICT) and at the same time ICT based technologies have altered the research dynamic.

In the first paper in this section, 'A Social, Psychological, Educational and Criminological Trials Register (SPECTR)', Petrosino *et al.* argue the case for systematic reviews in these fields on the Cochrane model. The issue of locating relevant studies is crucial for minimising bias in synthesising the evidence, and the register they describe provides a means of facilitating this. By locating over 10,000 reports of probable randomised trials in criminology, social policy, psychology and education they have made a valuable start to this process, though as they acknowledge, much more remains to be done.

In 'Should All Literature Reviews Be Systematic?' the authors (Badger *et al.*) keep the focus on the use of published material but argue convincingly that it is not the achievement of a systematic review that is the most critical feature but the exposure of the methodology to scrutiny. They point out some of the ways in which the process of carrying out a systematic review is more problematic than it might first appear. There is also a useful subplot about research students here.

In 'Randomised Trials or the Test of Time? The Story of Human Albumin Administration' (Roberts) we are shown an example of the results of one such review. This review showed that a widely used treatment for critically ill patients was in fact not justified by a synthesis of the best available evidence, and even that there were 'strong suggestions that it increases the risk of death'. The author illustrates some of the effects of publishing the review and draws attention to the improved possibilities for updating.

In a sense these papers are a reminder of what we already know – that there are many types of literature review. With the increase in the amount of research material available we need to deal with this in some way that is itself open to scrutiny. Perhaps now that many people have access to the web we may be more attentive to using the expertise of and listening to lessons from Information Specialists. Another thing to learn from researchers in science (and this one is easy) is their desirable habit of routinely naming the whole research team on publications.

## The Power of Indicators

The third section includes papers which focus on the use of indicator systems within the context of schooling. In this they get down to the nitty-gritty business of what happens at the interface between research and school improvement. In different ways they each explore problematic aspects of using research to improve schooling. They explore the day-to-day assumption that the more we know about performance and the more we understand about what contributes to achievement, then the more improvement will follow.

In 'Monitoring School Performance for "Standards-Based Reform"' Willms reviews findings linking research and socioeconomic factors. One of the many critical points raised in this paper is the question of variation at different levels in the system. Willms' analysis predicts that certain lines of intervention will be more likely to produce success than others. He identifies that there is more varia-

tion at classroom level than at school level, the implication of this being that the appropriate level of intervention is the teacher not the school. He further argues that closing poorly performing schools will dismiss effective as well as ineffective teachers.

Willms also addresses the issue of whether or not performance-based assessment increases student learning. He concludes that monitoring certainly drives schooling, but not necessarily in the direction wanted. He queries whether increased monitoring narrows the learning experiences offered to children and concludes that in at least some examples it has done.

In their paper, Visscher *et al.* report an attempt to follow up this last point. In 'Evidence on the Intended and Unintended Effects of Publishing School Performance Indicators' the authors attempt just that to detect the intended and unintended effects of publishing school performance indicators. In the event the authors found collecting the data was more difficult than had been expected and the evidence on the effects of publishing was problematic.

McWalters and Cheek report a large-scale attempt to improve schooling. In 'A State Accountability System as a Technology of Social Control: The Case of Rhode Island, USA', they argue that the technologies of state control embody values which necessarily shape the policy changes at each and every stage.

## Areas for Development

The papers in this special issue contain a number of significant ideas and developments. They include contributions to the improvement of research methodologies, to debates about the impact and presentation of research, and accounts of some exemplary studies. However, the most important lessons for researchers from these papers are arguably not in the detail of advances, although these are without doubt important. The most important lessons are about outward-looking research which creates a proper conversation with the citizenry and which maintains a dialogue with other branches of knowledge in order to do so.

## Correspondence

Any correspondence should be directed to Dr Robert Coe, Curriculum, Evaluation and Management Centre, Mounjoy Research Centre 4, University of Durham, Stockton Road, Durham DH1 3UZ, UK (r.j.coe@dur.ac.uk).

## References

- Fitz-Gibbon, C.T. (1996) *Monitoring Education: Indicators, Quality and Effectiveness*. London: Cassell.
- Sackett, D.L., Rosenberg, W., Gray, J.A.M., Haynes, R.B. and Richardson, W. (1996) Evidence-based medicine: What it is and what it isn't. *British Medical Journal* 312, 71–72.