

Research evidence, knowledge management and educational practice: lessons for all?

Paper for High-level Forum on Knowledge Management in Education and Learning, Oxford, March 18-19 2002.

Professor Ann Oakley
Director, EPPI-Centre and Social Science Research Unit
University of London Institute of Education

Terms such as 'knowledge management' and the 'knowledge economy/society' only loitered on the margins of my professional language until I was asked to participate in this seminar. My own education, which propelled me into a working lifetime in the Higher Education sector, was based on the unarticulated assumption that knowledge and educational institutions are intimately related in two ways: the education system is about the production and dissemination of knowledge, and what happens within it is itself knowledge-based. Teachers know how to teach and what they teach is worth knowing. But the nature of progress is very often that what we once took for granted turns out to be a mass of problems and complexities. Education is not 'simply' about the transfer of knowledge: it is about many other processes and outcomes as well. What is knowledge? Who is knowledge for? Who defines what we want to know about and why? How is knowledge used, and who benefits from (or is damaged by) these uses? Perhaps most problematic of all are the two questions: how does anyone know what they know? And, how do we get from knowledge to wisdom?

This paper discusses a small aspect of this troublingly complex picture: the evolution of the 'evidence-informed' education movement in England. It takes a synoptic look at the background and context in which this movement developed, describes the setting up of an initiative sponsored by the English Department for Education and Skills to promote the synthesis of research evidence, and discusses some of the early challenges experienced in taking this initiative forward. The paper argues that this practical experience in reviewing educational research is an important reference point for current debates about the application of the 'evidence movement approach' to educational research. It suggests that many criticisms of this approach are ill-founded in promoting centralized, systematic inquiry as the enemy of democratic knowledge, and in seeing the principles of systematic reviewing as necessarily at odds with methodological pluralism.

Swings and roundabouts: the 'science' of educational knowledge

The debate about whether education is an art or a science has gone on for a long time (Rusk, 1919). Central to this debate is an argument about the methods used to generate knowledge about effective and appropriate educational practices: theories about education, multiple understandings about the role and purpose of education, and teachers' individual 'expert' experiences, on the one hand, or systematically accumulated, replicable and generalisable evidence of the kind used to judge the therapeutic status of professional interventions in other fields, on the other? The other historical constancy is a degree of consensus about the chaotic and methodologically impoverished nature of much educational research (Eisner, 1979; Gage, 1972; OECD, 2000). 'Expert opinions, pooled judgements, brilliant intuitions, and shrewd hunches are frequently misleading,' wrote Julian Stanley in 1957. 'Ultimately they must be tested by the careful gathering of evaluative data if education is to advance on the basis of sound principles' (Stanley, 1957:198). 'Given the huge amounts of educational research conducted over the last fifty years or more,' remarked David Hargreaves in his lecture to the Teacher Training Agency in 1996, 'there are few areas which have yielded a corpus of research evidence regarded as scientifically sound and as a worthwhile resource to guide professional action' (Hargreaves, 1996:2).

Such opinions have acquired a new lease of life in the form of the 'evidence' movement. In the late 1970s and early 1980s a group of health service researchers in Oxford prepared the ground for evidence-based medicine by beginning a programme of systematic reviews on the effectiveness of health care interventions. The Cochrane Collaboration (<http://www.cochrane.org>) is now an international network of researchers, academics, practitioners and users committed to the principles of managing knowledge in such a way that it is quality assured, accessible, and cumulative. Cochrane reviews mostly, but not exclusively, focus on randomised experimental studies, and include a number of relevance to education, for example interprofessional education, and school-based driver education for the prevention of traffic crashes. The Cochrane Collaboration's sister organisation, the Campbell Collaboration (<http://campbell.gse.upenn.edu>), is adapting Cochrane methodology to prepare, maintain and disseminate systematic reviews of social interventions, but, being younger, is still working out its methods. The climate of thinking about public policy both in the USA and in the UK has shifted more firmly into the 'evidence' mode, with policy-makers declaring the need for a sounder science on which to base more 'joined up' decisions about policy issues (Cabinet Office, 1999). It has also become clear that there is a considerable forgotten history of scientific policy

experimentation and research synthesis (Oakley, 1998a). A return to this tradition is urgently needed, not just for policy-makers or even to improve education and other professional practices: it is the ordinary citizen who is potentially most disadvantaged by the lack of an open, systematic basis of evidence concerning the many interventions that intrude into every corner of life (Smith, 1996).

These developments make it more difficult to hide failures in a field of research activity which is of such critical public importance as education. Unlike health care, education is a compulsory intervention. Like health care professionals, teachers believe that they are only acting in the best interests of those at the receiving end of their work, and that it is their unique vocation to 'know' the best forms of practice. These parallels between education and health care have driven some recent criticisms of educational research, and have particularly drawn attention to the over-reliance in much educational practice on procedural 'craft' knowledge rather than more open research-based evidence (Hargreaves, 1996; Hillage et al., 1998; McIntyre and McIntyre, 1999; Tooley and Darby, 1998). Another way to put this is to describe the educational sector as dominated by non-codified but potentially codifiable knowledge (OECD, 2000:19). There are areas such as HIV/AIDS education where the health care and education systems have a shared responsibility to introduce a greater degree of codification into the knowledge base. The disadvantages of *not* doing this can, in such circumstances, literally be fatal. A recent project helping policy-makers, practitioners and researchers in southern Africa to develop a more scientific approach to HIV/AIDS education demonstrated how much of an unmet need there can be on the ground for support in moving towards a more systematic evidence-base (Stewart et al., 2001).

What is evidence?

Davies (1999:109) defines 'evidence-based education' as operating at two levels: the *use* of educational research and the *establishment* of sound research evidence through academic practices which attend to the criteria of scientific validity, high quality and practical relevance.

Systematic reviews are the primary method for managing knowledge in the evidence movement approach. This is because they synthesise the findings of many different research studies in a way which is explicit, transparent, replicable, accountable and (potentially) updateable. In these respects, systematic reviews differ from traditional literature reviews, which commonly focus on the *range* and *diversity* of primary

research using a selective, opportunistic and discursive approach to identifying and interpreting relevant literature (Badger et al., 2000; Davies, 2000). In traditional 'narrative' reviews, there is often no clear audit trail from primary research to the conclusions of the review, and important research may be missing, resulting in biased and misleading findings, and leading to puzzling discrepancies between the findings of different reviews.

Although there is a strong tradition of non-systematic reviews in social science, there are usually more of the systematic kind than might be expected. For example, a survey of effectiveness reviews in health promotion identified 398 reviews, of which 75 could be considered systematic (Peersman et al., 1999). There are over 300 systematic reviews in the Campbell Collaboration's Social, Psychological, Educational and Criminological Trials Register (Petrosino et al., 2000).

The EPPI Centre

Recognising the importance of systematic reviews in improving the evidence available to educational practitioners and policy-makers, the English Department for Education and Skills funded the Evidence for Policy and Practice Information and Coordinating Centre at the Social Science Research Unit, the University of London Institute of Education in 2000 to undertake a five year programme of work in this area. The main purposes of the funding are: to support groups of researchers, practitioners and users external to the EPPI Centre to carry out systematic reviews; to develop the tools, procedures and training needed to facilitate review group work; and create and sustain accessible web-based databases of systematically keyworded research literature, data from primary studies, and the results of systematic reviews. The EPPI Centre initiative is focusing initially on schools and students aged 0-16 years.

The EPPI Centre's DfES work is thus about managing rather than generating knowledge. It is based on a number of key principles about systematic reviews, some of which are common to systematic reviews in other topic areas, and some of which flow from the EPPI Centre's own previous experience of conducting systematic reviews of social interventions in areas such as health promotion and sex education. In the EPPI Centre approach, a systematic review:

- involves a series of explicit, discrete and standard stages;

- means specifying a particular, answerable research question, and criteria about what kinds of studies (by topic/ population group/ setting /research design etc) will be included and excluded in the domain of literature to be surveyed;
- requires clarity about which literatures will be searched for relevant studies, and how;
- includes making explicit, justifiable decisions about the methodological quality of studies regarded as generating reliable findings;
- needs some method of integrating the findings of individual, good quality studies;
- is credible only if it has involved input from research users at all stages of the review process; and
- is a much more time- and resource-intensive activity than a traditional literature review.

These principles are laid out and given operational flesh in a manual for review groups available as an online tool (<http://eppi.ioe.ac.uk>). The first part of the manual describes the work of the DfES initiative, the second goes through the methods of systematic reviews in education, and the third describes EPPI Centre procedures for coordinating and supporting the work of individual review groups. Also available on the same website is a document outlining a keywording strategy for classifying educational research studies, and references to some of the EPPI Centre's own systematic reviews which develop traditional review methodology in order better to meet the challenges of systematic reviews in social science. For example, a criterion for including studies in a review of workplace health promotion interventions was that interventions needed to have been developed in response to needs or views expressed by employees (Harden et al., 1999). A review of peer-delivered health promotion for young people included 'qualitative' studies of the processes involved in implementing different programmes (Harden et al., 2001). And a recently completed series of reviews on barriers to, and facilitators of, young people's mental health, healthy eating and physical activity set research on young people's views alongside evaluations of health promotion interventions to see what could be learnt about matches and mismatches between the policy and practice conclusions emerging from each type of research (Shepherd et al., 2001).

It is clear from these examples, and from the steps and procedures set out in the EPPI Centre training materials, that systematic reviews can include a wide range of

study designs. There is nothing intrinsic to the principles of systematic reviewing that specifies a hierarchy of evidence favouring prospective experimental studies with control groups. The process set up by the EPPI Centre to support review groups aims to be facilitative rather than authoritarian and to engage in mutual learning about the best ways to establish a collaborative approach to research synthesis in education. There is no 'central directive' about prioritising certain kinds of review question or certain kinds of research design.

Ten review groups have been established in the first two years of the EPPI Centre education initiative in the following areas: English teaching; assessment and learning; school leadership; gender and education; post-compulsory education; inclusive education; early years; thinking skills approaches to effective learning and teaching; modern languages; and the impact of continuing professional development on classroom teaching and learning. The methods for facilitating the review work of these groups were piloted in a systematic review of strategies to support pupils with emotional and behavioural difficulties (EBD) in mainstream primary classrooms conducted by two researchers at the Institute of Education and the National Foundation for Educational Research (Evans and Benefield, 2001).

EPPI Centre review groups go through a formal registration process which includes peer refereeing of plans and protocols. Each review undertaken by a review group requires a detailed review protocol which defines the review question, inclusion and exclusion criteria and strategies for searching the literature. All protocols are placed on the EPPI Centre website for open comment. Most reviews are done in two stages: *a mapping stage*, in which relevant literature is captured and systematically keyworded to provide a descriptive account of the research effort in that particular area; and *an in-depth review stage*, in which a subset of the literature found is examined and interrogated in more detail and data extracted from primary studies. An important feature of the reviewing software developed by the EPPI Centre is that it allows data to be entered on a range of study designs. These data are then analysed together for individual reviews, but they can be combined differently to answer other review questions; reviews can also easily be updated by adding data from new primary studies.

Each review group is scheduled to produce one review per year and the reviews will be available on line in a database called REEL (Research Evidence in Education

Library. The first reviews are now coming in and we expect them to be available on REEL by summer 2002.

The first reviews

The first reviews from the six review groups registered in the first year show considerable diversity in terms of the questions asked and the kinds of studies included in the review. The review questions are: 'what is the impact of networked ICT on literacy learning in English for pupils aged 5 to 16?' (the English RG); 'what is the impact of summative assessment on pupils' motivation for learning?' (the Assessment RG); 'what is the impact of school leadership and management on student/pupil outcomes?' (the Leadership RG); 'what kind of strategies improve equal opportunities in gender for pupils in mixed sex primary schools?' (the Gender RG); 'what is the impact of the financial circumstances of learners on their engagement with learning?' (the Post-compulsory RG); and, 'what evidence is there that mainstream schools can act in ways which enable them to respond to diversity so as to facilitate participation by all students in the cultures, curricula and communities of those schools?' (the Inclusive Education RG).

Types of research given prominence in the review protocols are as follows: outcome evaluations, process evaluations and descriptive studies (the English RG); intervention and non-intervention studies (the Assessment RG); intervention studies and school effects and improvement research (the Leadership RG); outcome evaluations and process evaluations (the Gender RG); studies 'which seek to understand the links between financial circumstances and engagement' (the Post-compulsory RG); ethnographic case studies, outcome and process evaluations, survey research and other types of case study (the Inclusive Education RG).

Challenges of the review process

This process of managing knowledge about educational practices through central support for review group work is itself an experiment which will need careful evaluation. The first two years' work have felt much like stumbling hand-in-hand through ill-lit passageways inspired by a dim light at the end of the tunnel. Sometimes there have been minor collisions in the shadows, but there have also been moments of collective illumination. Ultimately, the proof of the pudding will be in the eating – whether the reviews produced in the form proposed are regarded as valuable and are in fact used by the various stakeholders in educational research.

Some lessons have emerged in these early experiences of this approach to managing educational research knowledge. The list below draws on the pilot EBD review, RGs' reports of their first year's work, and observations of the training workshops and other EPPI Centre contacts with RGs. The challenges can be divided into technical, and conceptual/intellectual:

Technical challenges

1. *Lack of experience of the skills and procedures* (eg searching, classifying literature, keywording, data extraction) needed in systematic reviewing. Very little of the research training currently on offer in the social sciences or in teacher education includes components which prepare people for systematic review work. This is a problem for the core academic members of a RG, but also for the involvement of teachers and other educational research users.
2. *Reliance on relatively 'undeveloped' electronic databases*, many of which (eg ERIC and PsycLIT) lack comprehensive keywording, a thesaurus of standardized search terms and sophisticated search strategies, and each of which require individualized approaches. This highlights the importance of handsearching, which is a very time-consuming exercise. For example, the Assessment and Learning RG spent a frustrating four months on searching 63 journals electronically and by hand with a much greater use of handsearching than initially anticipated, largely because of the inadequacies of electronic databases.
3. *The relatively low yield of usable studies* derived from the searching process. For example, the English RG identified 1871 potentially relevant papers, but excluded 1510 of these after screening the title/abstract, and a further 131 after retrieving and reading the whole paper. This experience of having to search haystacks to find needles is common in systematic reviews (Oakley and Fullerton, 1996).
4. *The length of time systematic reviews take*. Searching for literature is a lengthy business, scrutinizing titles and abstracts for relevant information is often exhausting and disappointing, retrieval of hard copies of studies for indepth review is another time-consuming task, and extracting data from individual studies for the review can take 3-5 person hours per study. This means that 'such reviews should be given the status and time allocation of a significant research project' (Evans and Benefield, 2001:539). It also means that the real cost of producing reviews is significant. The DfES, through the EPPI Centre, provides some pump-priming money for the RGs (£20k for the first year, £15k thereafter), but experience of the first RGs suggests that the

real cost per review is likely to be closer to £75k; the in-house reviews the EPPI-Centre does for other funders run from £50-£80k per review, and detailed reviews which tackle important methodological questions at the same time can cost in excess of £100k. Given these resource demands, pump-priming money can be rapidly used up, resulting in less-than-ideal shortcuts in the review process.

Conceptual/intellectual

5. *Defining the initial research question* for a systematic review. This can be surprisingly tricky. Groups may find their collective expertise in an area not necessarily an aid to clear thinking about how to identify a single review question and define the key terms included in such a question to the satisfaction of all. For example, the Inclusive Education RG had problems defining the terms 'inclusive' and 'organisational characteristics' in its initial review question in such a way that these were acceptable to the different ideological positions of RG members (Torgerson et al., in press). ('Inclusive' became the promotion of 'participation of all students in the culture, curricula and communities' of their schools, and 'organisational characteristics' became 'ways in which schools can act'.) Further definition may be necessary for the question asked in the indepth stage of the review. For instance, the Gender RG's indepth review question narrowed the scope of its initial review question to become 'what kind of strategies delivered by teachers or researchers in the classroom reduce stereotypical gender constructions among girls and boys in UK mixed sex primary schools?' The final protocol for a review can be a lengthy task: the English RG's review protocol went through 6 drafts, for example. One reason why specifying review questions and criteria can be such a lengthy, iterative process is because such discussions are precisely not just technical; they provide a forum for debate about important theoretical and ideological issues. Systematic reviews are in this sense as much a theory-based as a pragmatic form of inquiry.
6. *Deciding how to define 'sound' studies* (of whatever research design). While some systematic reviews may include a narrative synthesis of all the located studies, it is more usual to filter out for detailed discussion/analysis a subset of the more methodologically robust studies. However, the notion that some research studies may be 'better' methodologically than others can offend the egalitarianism of some systematic reviewers. There can be obvious problems in defining some studies as less 'good' than others when the field is relatively small, and researchers all know one another. Even when the need for quality

criteria is agreed, it can be difficult to settle on particular criteria for identifying good quality studies. It can also be very difficult to decide when such criteria have been met, since a surprising proportion of studies are very inexplicit, not to say vague, about the designs they have used.

7. *Operationalising 'user involvement'*, a key principle of the EPPI Centre approach to systematic reviews. The boundaries between 'users' and 'producers' of research are often fluid. For example, academics are both research-producers and research-users. Representatives of some kinds of 'users' (eg members of teaching organisations) are easier to find than others (eg primary school students). While all the EPPI Centre RGs involve some users, it has been easier to identify practising teachers than representative parents or students interested in systematic review work.

Managing knowledge: lessons for all?

A recent report of the National Research Council in the USA on *Scientific Inquiry in Education* describes the accumulation of scientific knowledge over time as 'circuitous and indirect', and notes that, 'It often traverses highly contested territory...'

(Shavelson and Towne, 2001). The report notes six fundamental scientific principles that apply in the field of educational research: posing significant questions that can be answered empirically; linking research to relevant theory; using methods that permit direct investigation of the question; providing a coherent and explicit chain of reasoning; replicating and generalizing across studies; and disclosing research to professional scrutiny and critique.

While few would dispute the general importance of these principles, a sizeable voice in the educational research field today strongly contests what is seen as the unwarranted transfer to the educational domain of a 'positivistic' model of knowledge dominated by a 'hierarchy of evidence', limited questions about 'what works', outdated notions about the role of 'procedural objectivity' and a disregard for the tenets of postmodernism which question the validity of any *a priori* framework of enquiry (Atkinson, 2000; Elliott, 2001; Hammersley, 2001). Perhaps most worrying of all is the accusation that a coordinated structure for preparing systematic reviews in education reeks of unwarranted monitoring and control (Ball, 2001).

These criticisms from the academic research and practitioner communities of the evidence movement approach to knowledge management are not confined to education. They repeat themes of earlier attacks on evidence-based medicine and

evidence-based health promotion: debates in the latter arena, particularly, have much in common with current educational critiques, especially concerning the prioritisation of 'qualitative' forms of enquiry and the 'evidence' of expert opinion (Oakley, 1998b).

Paradigm 'wars' setting systematic reviews as a way of managing knowledge against other approaches to knowledge production signal a 'community that is not at ease with itself', and one in which the mutual respect underlying the often-proclaimed commitment to pluralism is often strangely lacking (Hargreaves, 2001:2). A level of disquiet about the shift in thinking about knowledge and policy towards a more systematic and open system is understandable, especially given the role of government funding in supporting the development of the relevant infrastructures. But many of the anxieties currently being voiced are based on misunderstandings of what research synthesis actually in practice means: whose interests it is for, who will produce and use it, and who will benefit from it. Some of the arguments of the 'anti-evidence' movement sit oddly beside the attempt to open up and systematise knowledge about educational practices which is described in this paper. These experiences speak to a process of collaborative, open and democratic working directed to the endpoint of a freely available updateable library of reviews of research evidence. This is likely to inject much needed science *and* commonsense into what we as 'a knowledge society' collectively 'know': it may even, eventually, take us on that other dimly lit path from knowledge to wisdom.

References

- Atkinson E (2000) In defence of ideas, or why 'what works' is not enough. *British Journal of Sociology of Education*, 21(3):317-30.
- Badger D, Nursten J, Williams P, Woodward M (2000) Should all literature reviews be systematic? *Evaluation and Research in Education*, 14(3&4):220-30.
- Ball SJ (2001) 'You've been NERFed!' Dumbing down the academy: National Educational Research Forum: 'a national strategy – consultation paper': a brief and bilious response. *Journal of Education Policy*, 16(3):265-68.
- Cabinet Office (1999) *Professional Policy-making for the 21st Century*. London: Stationery Office.
- Davies P (1999) What is evidence-based education? *British Journal of Educational Studies*, 47(2):108-21.
- Davies P (2000) The relevance of systematic reviews to educational policy and practice. *Oxford Review of Education*, 26 (3&4):365-78.
- Eisner EW (1979) *The Educational Imagination: On the design and evaluation of school programs*. New York: Macmillan.
- Elliott J (2001) Making evidence-based practice educational. *British Educational Research Journal*, 27(5):555-574.
- Evans J, Benefield P (2001) Systematic reviews of educational research: does the medical model fit? *British Educational Research Journal*, 27 (5):527-41.
- Gage NL (1972) *Teacher Effectiveness and Teacher Education: The search for a scientific basis*. Palo Alto, Calif.: Pacific Books.
- Hammersley M (2001) On systematic reviews of research literatures: a 'narrative' response to Evans and Benefield. *British Educational Research Journal*, 27(5):543-54.
- Harden A, Peersman G, Oliver S, Mauthner M, Oakley A (1999a) A systematic review of the effectiveness of health promotion interventions in the workplace. *Occupational Medicine*, 49(8):540-48.
- Harden A, Peersman G, Oakley A (1999b) Identifying primary research on electronic databases to inform decision-making in health promotion: the case of sexual health promotion. *Health Education Journal*, 58:290-301.
- Hargreaves D (1996) Teaching as a research-based profession: possibilities and prospects. Teacher Training Agency Annual Lecture. London: TTA.
- Hargreaves D (2001) The Nuttall Memorial/Carfax Lecture. British Educational Research Association Annual Conference.
- Hillage J, Pearson R, Anderson A, Tamkin P (1998) *Excellence in Research in Schools*. London: Department for Education and Employment/Institute of Employment Studies.
- McIntyre D, McIntyre A (1999) *Capacity for Research into teaching and Learning*. Final Report to ESRC Teaching and Learning Programme. Cambridge: School of Education, University of Cambridge.
- Oakley A (1998a) Public policy experimentation: lessons from America. *Policy Studies*, 19 (2):93-114.
- Oakley A (1998b) Experimentation in social science: the case of health promotion. *Social Sciences in Health*, 4(2):73-89.
- Oakley A, Fullerton D (1986) The lamp-post of research: support or illumination? In Oakley A, Roberts H (eds) *Evaluating Social Interventions*. Ilford: Barnardos.
- OECD (2000) *Knowledge Management in the Learning Society*. Centre for Educational Research and Innovation.
- Oliver S, Peersman G, Harden A, Oakley A (1999) Discrepancies in findings from effectiveness reviews: the case of health promotion for older people in accident and injury prevention. *Health Education Journal*, 58:66-77.

- Peersman G, Harden A, Oliver S, Oakley A (1999) *Effectiveness Reviews in Health Promotion*. London: EPI Centre.
- Petrosino A, Boruch RF, Rounding C, McDonald S, Chalmers I (2000) The Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR) to facilitate the preparation and maintenance of systematic reviews of social and educational interventions. *Evaluation and Research in Education*, 14: 3&4, 206-19.
- Rusk RF (1919) *Experimental Education*. London: Longmans, Green and Co.
- Shepherd J, Garcia J, Oliver S, Harden A, Rees R, Brunton G, Oakley A (2001) *Barriers to, and Facilitators of, the Health of Young People: A systematic review of evidence on young people's views and on interventions in mental health, physical activity and healthy eating*. Volume 1, Overview; Volume 2 Complete report. London: EPPI Centre.
- Shavelson RJ, Towne L (eds) (2001) *Scientific Inquiry in Education*. Washington, DC.: National Academy Press.
- Smith AFM (1996) Mad cows and ecstasy: chance and choice in an evidence-based society. *Journal of the Royal Statistical Society*, 159(3):367-83.
- Stanley JC (1957) Controlled experimentation in the classroom. *Journal of Experimental Education*, 25:195-201.
- Stewart R (ed.) (2001) *The HIVSA Project Report*. London: Social Science Research Unit.
- Tooley J, Darby D (1998) *Educational Research – a Critique*. London: Office for Standards in Education.
- Torgerson C, Roberts B, Thomas J, Dyson A, Elbourne D (in press) Developing protocols for systematic reviews in education: early experiences from EPPI Centre review groups. *Evaluation and Research in Education*.