High quality reporting of research: lessons learned from systematic reviewing in education

Judy Sebba and Diana Elbourne

Context

- 1998 review of educational research in England concluded that:
  - Educational research had little impact on policy or practice
  - It was characterised by small scale, fragmented studies and a lack of coherence across funders
- OECD review in 2002 noted progress but identified need for further training in reviewing

Major initiatives to address findings

- National Educational Research Forum (& Funders' Forum)
- Evidence for policy and Practice (Eppi) Centre to develop systematic reviews
- Greater investment in research centres & longitudinal studies
- National Teacher Research Panel
- Initiatives to improve coherence & dissemination e.g. websites of current research, re-written journal articles, etc.

Initiatives were aimed at:

- Better use of existing evidence by policy-makers, practitioners, users & researchers
- Investment in better future knowledge base

Eppi reviews

- 16 established review groups
- 9 new groups addressing research on teacher education
- 2 being undertaken collaboratively – one with Colorado State University on transition
- Ongoing reviews undertaken directly
- 15 published reviews in education & further 10 likely by Dec 03

Criteria in data extraction in Eppi reviews

- Undertaken by 2 reviewers – software helps identify agreement/disagreement
- Study aims and rationale
- Research questions
- Methods including design, information about participants, sampling strategies, setting, data collection & analysis
- Findings and conclusions

Judging how much weight to give studies

- Eppi reviews draw on studies utilising the full range of methodologies (unlike other systems)
- Based on premise of fitness for purpose
- Judgement about weight of evidence (WOE) each study contributes based on:

General quality e.g. completeness of data, use of methods appropriate to study question, conclusions are derived from findings (quality criteria explicitly stated by RGs)

Specific to review e.g. appropriateness of study design to review questions, relevance of focus to review questions

- Overall assessment of WOE crudely as high/medium/low

The Research Informed Practice Site

www.standards.dfes.gov.uk/research

Provides user-friendly, interactive research digests for users, including teachers, governors and parents. Current themes include:

- Thinking skills
- Assessment for learning
- Early years
- Literacy and numeracy
- Further Education

contents:

Effects of a Cognitive Acceleration Programme on Year 1 pupils

Authors:
Philip Adey, King's College London
Anne Robertson, London Borough of Hammersmith and Fulham
Grady Venville, King's College London

Publisher:
British Journal of Educational Psychology (2002), Vol. 72, pp. 1-25

Introduction:

Can teachers actually improve the thinking skills of young children?

Many recent studies have shown that teachers can affect the way in which their pupils' thinking develops. Philip Adey, Michael Shayer and other researchers, in earlier cognitive intervention programmes, had aimed to accelerate the development of thinking processes for children in Years 7 and 8. In the study

Publisher:
British Journal of Educational Psychology (2002), Vol. 72, pp. 1-25

Introduction:

Can teachers actually improve the thinking skills of young children?

Many recent studies have shown that teachers can affect the way in which their pupils' thinking develops. Philip Adey, Michael Shayer and other researchers, in earlier cognitive intervention programmes, had aimed to accelerate the development of thinking processes for children in Years 7 and 8. In the study summarised here the authors wanted to find out if a similar approach could be taken with children as young as 5 and 6. The authors designed and investigated the effect of a cognitive intervention programme on the development of thinking skills of 338 children in Year 1 in 10 primary schools in a socially deprived area of the London borough of Fulham and Hammersmith. The research was supported by the LEA in that area. The findings showed that such a programme could indeed be successful. The authors hope that the programme will bring beneficial long-term effects to pupils' academic development which in turn will influence positively their social development, and eventually their employment prospects. Whilst this paper reported initial findings, further reports are planned.

What have we learned 1

From findings of first reviews:

- Few were judged to provide high overall WOE for the review
- Few provided high quality WOE based on quality of the study itself – but is this due to quality of research or of reporting?
- Review process has capacity to both:

➢ Raise quality of future primary studies

➢ Raise future quality of reporting

What have we learned 2

(drawing on Bell et al, 2002 and review reports)

- Electronic abstracts less informative than h.copy
- 36% of the 97 journals used in TRIPs (as above) work do not provide abstract electronically
- ERIC provides list of criteria for abstractors include: subject matter, scope, purpose, intended audience, intended use, results, etc., but not sample or context characteristics
- ERIC abstracts often have key info missing
- Original author's abstract reduced by 50-70% in ERIC

What have we learned 3

(drawing on Bell et al, 2002 & Oakley, 2003)

- Electronic databases are inadequate – lack of comprehensive key-wording.
- Titles are misleading – in particular, clever titles aimed to attract or those relating to previous ones e.g. *The Protean Spirit of Jeff Lewis* (Oxford Review of Education)
- Abstracts are indicative rather than informative e.g. often do not include characteristics of students or schools
- Reviewing is therefore additionally expensive & laborious

Characteristics of abstracts
from Bell et al, 2002 BERA paper

Analysed 55 abstracts from 4 ‘reputable’ journals:

- >half were informative, included main criteria
- Most were indicative often covered purpose, methods or scope but not sample, findings
- 2:3 some details of sample but only 1:3 size
- Many read like an introduction rather than summary
- Many had key info missing especially dates

Example 1: Cambridge Journal of Education, 2002, 32, 1

Teachers and learners use spoken language in the three part mathematics lesson advocated by the National Numeracy Strategy (NNS). This document recognises the importance of language by emphasising 'the correct use of mathematical vocabulary' in the raising of standards. Research into the use of vocabulary in science suggests that the use of scientific words does not necessarily demonstrate conceptual understanding. Pupils and teachers appear to ascribe different meanings to scientific words because of their relative positions on the novice/expert continuum. To explore whether mathematical words could cause similar problems, data collected from six groups of teachers and learners was subjected to discourse analysis to provide evidence of how spoken language was used and how meaning and understanding were achieved. The implications of the findings on the use of language by both teachers and learners in the NNS mathematics lesson, including vital areas for reflection by teachers, are fully considered.

Example 2: Educational Evaluation and Policy Analysis, 2002, 24, 3

State and school districts are looking for policies they believe will improve student performance. As a result, assessments have proliferated, stakes have increased, and specific curriculum and instructional approaches are being mandated. In this study, we probed beneath students' failing scores on a state reading assessment to investigate the needs of struggling students and implications for policy. We found that scores on state tests mask distinctive and multifaceted patterns of students' reading abilities that require dramatically different instructional emphases. We explore the implications of this complexity for state and local reform efforts that target improved teaching and learning

Example 3: Educational Evaluation and Policy Analysis, 2003, 25, 1

This article explores patterns of special education services during the elementary grades among children who participated in either the Child-Parent Center (CPC) Preschool Program or other early childhood programs in the Chicago Public Schools. The study sample included 1,377 low-income, racial minority children in the Chicago Longitudinal Study. Controlling for family background characteristics that might affect educational performance, children who participated in CPC Preschool had a significantly lower rate of special education placement (12.5%) than the comparison group (18.4%), who participated in an alternative all-day kindergarten program. The estimated impact of CPC preschool intervention was best explained by the cognitive advantage hypothesis. This article provides support for the long-term impact of the CPC preschool intervention on special education outcomes.

Some factors limiting quality

- Lack of recognition of importance of high quality reporting for academic, policy and practice readers as well as for reviewers
- Lack of training of editors, peer reviewers or researchers
- Academic editing is unrecognised and isolated
- Pressures to get published due to link with funding has led to increase in no of journals and delay in publication
- Not related to type of research – see problems in medicine
- Poor quality reporting often but not always reflects poor quality research

Ways Forward 1

- Training in reviewing develops writing/abstracting skills
- Training for editors and peer reviewers
- Greater recognition of editing and reviewing in systems for external assessment of research – to increase incentives
- Increase use of structured abstracts – Hartley (2003) reports preferred as provide better information and standardise presentation so search & retrieval much easier
- Develop international set of standards for abstracts (education? or social science?)- consistency v. straightjacket e.g. Haynes, Hartley, ERIC, etc
- Develop international association of education (or social science?) editors (Squires, 1995 describes medical one)

Examples of criteria for abstracts

- Haynes: Objective/research question, study design, setting, participants, intervention, main outcome measures, results, conclusions
- ERIC (author manual): Subject matter, scope, purpose, publication type, audience, relationship to other work, intended use, results/findings
- Hartley: Abstract evaluation checklist includes previous research, aims/purposes, participants (number, origins, gender, age, allocation to groups), measures, results, conclusions, limitations, further research

The APA publication manual (2001) states that a well-prepared abstract needs to be dense with information but also readable, well organized, brief and self-contained.

Can we succeed in getting the balance between brevity and informative and between consistency and rigidity?

References

Bell, M., Cordingley, P., Curtis, A., Evans, D., Hughes, S. & Shreeve, A. (2002) Bringing research resources to practitioner users via web technology: lessons learned to date. Paper to BERA Annual conference, Sept, 2002

Hartley, J. (2003) Improving the clarity of journal abstracts in Psychology: The case for structure. *Science Communication*, 24, 3, 366-379

Haynes, R.B., Mulrow, C.D., Huth, E.J., Altman, D.G.& Gardner, M.J. (1990) More informative abstracts revisited. *Annals of Internal Medicine*, 113, 69-76

Oakley, A. (2003) Research evidence, knowledge management and educational practice: early lessons from a systematic approach. *London Review of Education,* 1, 121-33

Squires, B.P. (1995) A global network for medical journal editors *Canadian Medical Association Journal,* 152, 11, 1757-1759