'Hard health' and 'soft schools': research designs to evaluate SLT work in schools

Elspeth McCartney
Speech and Language Therapy Department,
University of Strathclyde, Glasgow, UK

Abstract

While systems approaches are useful for evaluating speech and language therapists' (SLT) work in individual school contexts, there is a need to undertake studies detailing in a replicable format the interventions offered to children and for studies at all levels to assess whether these interventions work, using validated scientific techniques. There is a demand for such studies to meet the National Health Service objective of using evidence-based approaches, which offer the best interventions available. Education researchers are being asked to address similar issues, and an overview is given of the type and levels of research used in the two sectors. It is suggested that health and education research are moving closer together and that SLTs in schools should undertake exploratory, group and cohort studies to further develop effective therapies.

Introduction

Writing in this journal (McCartney and van der Gaag, 1996, see also McCartney, 1999) on the issue of how speech and language therapists (SLTs) might evaluate their work in schools, I argued that evaluation techniques commonly used by education services were more useful than those of the health service. The use of a systems model was outlined that captured some of the complexities of SLT service delivery. Such a model can ask and answer questions about whether services have developed aims which foster collaboration; whether they have developed structures to make and record decisions jointly; how a child moves into, through and out of services;

Address for correspondence: Elspeth McCartney, Speech and Language Therapy Department, University of Strathclyde, Southbrae Drive, Glasgow G13 1PP, UK. E-mail: e.mccartney@strath.ac.uk

and whether the opinions of parents and others are sought. There was also recognition that many SLT interventions had a 'habilitative' care aim, where it was hoped and intended that children would progress in aspects of language and communication. Children's progress fitted into the model in terms of their attainment of individual education programme (IEP) targets, used both as a measure of an individual child's progress to guide the planning of therapy and aggregated to form a summary account of how a service was fulfilling its overall aim of helping all children. This way of summarizing attainments is comparable to the self-evaluation approaches used in schools to assess learning and teaching, which use IEP targets for children with additional learning needs to encapsulate attainment (SOEID, 1999; HMIE, 2002). These approaches accept that successful interventions depend on numerous interacting factors, that measures of progress should be child-specific, and that progress can be measured successfully by the 'bendy rulers' of IEP goals and targets.

However, while schools have developed the process of evaluation along such lines, health services have moved towards different models of demonstrating efficacy. In line with a major NHS initiative to deliver evidence-based medicine, that aims to ensure clients receive the most efficacious interventions available, health services have been encouraged to use 'harder' research paradigms to show that what they do works. The aim is to develop interventions that are not context-specific, but which are effective in many real-life contexts and with many professionals and children in ways that are comparable to medical treatments. This search for evidence to support practice applies to SLTs in the UK, where they are mostly employed by health services, and so includes SLTs' work in schools. Indeed, as demand for SLT services in schools increases there are concomitant requirements to show that such input is worth having and that children can benefit in specified ways from the services provided. The drive for evidence-based approaches that are not context-dependent is becoming influential and this paper argues that SLT services must now engage with the issue and find ways of presenting evidence of effectiveness that are acceptable to their NHS employers, as well as to the schools in which they work.

'Hard health' evidence

The level of evidence of effectiveness required in NHS settings is high and depends upon controlling the factors that influence the outcomes of therapy interventions, in this case, benefits to the child. There are recognized types of

research design that control more or less thoroughly for the many factors known to influence intervention outcomes. In order of most to least robust, there are:

- randomized control trials (RCTs);
- controlled cohort studies;
- case study series;
- individual case studies
  (Centre for Evidence-Based Medicine 2003).

Below the level of these studies is 'professional opinion', ranked lowest of all.

Randomized control trials

Because it controls very well for relevant variables, the RCT is considered the most powerful research design, the gold standard against which other designs are measured. In RCT designs the number of children needed to detect an intervention effect is determined before the study takes place, children are allocated to interventions on a random basis, interventions are compared to other therapies or to current standard practice and assessment of progress is made on predetermined outcome measures. Neither the child nor the SLT can remain unaware that therapy is being delivered, but post-intervention measurement is carried out by assessors blind to the intervention received, to avoid bias. This research design takes away professional decisions about which child gets which type of intervention, so the interventions must be equally plausible in their predicted outcomes (a feature known as 'equipoise' amongst interventions). The control achieved over relevant variables makes the RCT the research design of choice in the NHS, and it is 'considered by many the *sine qua non* when addressing questions regarding therapeutic efficacy, whereas other study designs are appropriate for addressing other types of questions' (Clarke and Oxman, 2003: 4.2.4).

Jones *et al.* (2001) give a detailed account of the factors to consider when setting up an RCT with an example from therapy for early stammering, but there are very few school-based examples using this research design. This is partly because RCTs are difficult and expensive to organize in school settings. It is quite unusual to be able to allocate school children at random to interventions, especially children with a record or statement of special educational needs. Language difficulties are also heterogeneous and within any school class children may require quite different intervention techniques. If we predict that the effects of SLT intervention will be small and hard to detect amongst the rich experiences provided by schools, we need large numbers of children in any study, which can be difficult for small services

to arrange. We also need careful specification of what therapy interventions consist of, and practitioners must stick to their therapy plans. This aspect can be expensive to set up and monitor and may introduce an unacceptable level of coercion into the therapy process. Assessment of outcome is best done on standardized measures, but these may not exist. It should be carried out by assessors blind to what intervention a child received and these people must be appropriately qualified, available, contracted and paid.

Such real organizational and cost limitations mean that RCTs are seldom attempted in schools, although there is at present one large-scale RCT ongoing (Boyle *et al.*, 1999) looking at direct and indirect, group and individual therapies for school-aged children with primary language impairment. This study has constructed a therapy manual, collating published approaches to comprehension monitoring, vocabulary development, later grammar and narrative and linking these to therapy activities and materials (McCartney *et al.*, 2004). The use of a manual provides flexibility, but also allows replicable language therapy for research purposes. If its usefulness is validated by the research children's outcome measures, it can be used by other researchers to replicate results and test whether this intervention can generalize to other contexts.

Cohort studies
The next level of evidence is provided by cohort studies (Sackett *et al.*, 2000). Here a complete group of children is treated and followed up over a period of time. If every child in the cohort is assessed and the reasons for their inclusion in the study are clear, some sensible deductions can be made about the effectiveness of therapy. Measuring another highly similar group whose intervention varies in kind or in amount to serve as a control further strengthens cohort studies. A major difference between RCTs and cohort studies is that the decision about who receives intervention is not usually randomized, but decided by the relevant professional. This means that 'intervention' children might differ from 'control' children on factors that can affect outcome (called 'confounders'). Where these confounding variables can be predicted (for example age, severity or type of language difficulty) they can be carefully documented, to check that the two groups do not differ on such variables or to adjust for them when considering outcomes. The obvious limiting factor is that professionals can only deal with confounders that are already known and have been measured (Sackett *et al.*, 2000: 157–58): unpredictable factors can remain as undetected influences on results. This is clearly bad science and provides misleading information on what works and why.

Nonetheless, a cohort study may show an overall therapy effect and is better than no comparison. Cohort studies are strengthened by providing clear information about the children being studied and what happened to them, with all children's results reported. Studies still need to report on a sufficiently large number of children (where 'sufficient' is a statistical concept that has to be estimated before the study starts) and the use of standardized measures of outcome makes results easier to interpret. Again, it is important to know just what therapy comprised, so that it can be replicated with future groups and blind assessment is needed. 'Before intervention', 'after intervention' and 'follow-up' measures are still needed, and if several base-line measures are collected over time this can add to our certainty that intervention had an effect. Gersten *et al.* (2000) give a helpful technical account of the factors to be considered when conducting such studies relating to special education but equally applicable to therapies. As an example of this approach, a cohort study (McCartney *et al.*, 2003) using therapy procedures from the Boyle *et al.* (1999) project, but with a different model of service delivery is currently being carried out.

Case study series
A less powerful design, often used when starting to develop new interventions, is the case series where a number of children with similar, specified, needs undertake the same intervention, with outcomes reported. These studies are not controlled, so it is not possible to know what might have happened if no intervention had been offered, or if a different intervention approach had been taken. Nonetheless they do show some preliminary evidence that the therapy package moved the children forward and if similar interventions show comparable effects on many children, the evidence for efficacy becomes more convincing. The difference between this design and the 'IEP targets achieved' counts used in schools is that the intervention would be specified in detail, and would remain as similar as possible across children, since the aim is to assess whether the intervention procedure works with different children. Some common outcome would also be measured. Given the potential useful-ness of this approach, it is surprisingly underused by SLTs in schools.

Individual case studies
Turning to the remaining levels of evidence, case studies of individuals can offer depth and detail and are useful for sharing ideas in the early stages of developing intervention approaches. They must of course later be replicated if they are to provide plausible evidence of a general effect. Professional opinion may be accurate and is a good place to start looking for effective practice, but

if not backed up by stronger evidence it is unlikely to persuade others of its validity.

The need for more evidence

As the NHS is keen to deliver effective interventions, high-level evidence is being sought to help make decisions about priorities and which treatments will be funded. Unfortunately for SLTs who wish to show effective work in schools, most of the current evidence, at least for language impairment, is at the level of professional opinion. A search of the four relevant academic data-bases (*Medline, PsyInfo, ERIC, LLBA*) for studies of treatment efficacy in children with language impairment aged 6–11 years using search techniques based on those of Law *et al.* (1998) revealed only four controlled studies of intervention; an additional unpublished research study was found via other sources (McCartney *et al.*, 2004). Only one older study (Methany and Panagos, 1978) was an RCT, the other four were controlled cohort studies involving small numbers of children. Uncontrolled case studies appear in the literature from time to time and with further systematic reviews being undertaken more work may come to light. However, there is at present an overwhelming lack of evidence about intervention techniques and their effectiveness.

The lack of good evidence for effectiveness can limit the work of the SLT in various ways. It can be hard to gain acceptance by NHS colleagues that SLT work in schools has value, and this can lead to difficulties in maintaining service levels and in funding new developments in competition with other NHS priorities. There are less dramatic but still relevant problems when developing research programmes, as there appear to be no validated 'treatments' to investigate. There is restriction on SLT pre- and in-service education and on the development of the profession, since good practice cannot be transmitted rapidly using research evidence. Being a profession without much evidence of effectiveness is bad for morale within the 'hard-evidence'-based NHS. More importantly, the lack of research leaves a large number of unanswered questions about whether children in school are getting an excellent or even adequate service.

'Soft schools' evidence

The research climate in education has developed somewhat differently. Perhaps surprisingly to NHS colleagues, there has been remarkably little

attempt in the UK to evaluate teaching methods or programmes in the way in which medicine evaluates treatments. At present most UK educational research relies on small-scale, qualitative studies and there is a lack of expertise in large-scale numeric studies, especially field trials derived from laboratory experimental designs (Gorard, 2002: 2). RCTs are very rare, although Torgerson and Torgerson (2001) recommend their (re-) introduction, citing some large-scale educational innovations that might have been more accurately evaluated by such methods (they also give useful information about how to set up randomized trials in an educational context).

There are practical and philosophical (and doubtless political) reasons for this state of affairs. At a practical level, small-scale studies tend to be cheap to run and can be carried out by teacher-researchers. At a philosophical level there is resistance amongst a number of educational researchers to some evidence-based approaches, on the grounds that educational practice may not be open to 'objective' assessment of effectiveness, but depends rather upon value judgements about desirable ends and appropriate means (Hammersley, 2001). Context-specific studies have predominated because, as noted above, children are not to be detached from their educational context and investigating and describing what happens in a particular context may be more useful to a service than attempting to find context-neutral teaching approaches. Context-specific studies are likely to make sense to practitioners and answer some of the questions asked by teachers about classroom practice, just as they answer SLTs' questions about the effectiveness of a particular type of service delivery. Methodologies that ensure rigour in constructing and reporting upon contextualized, flexible studies have been developed and a clear review of these issues is presented by Robson (2002).

However, the lack of studies that allow generalization across contexts, and the lack of hard evidence to inform educational practice has been criticized. Such criticism is perhaps most clearly articulated in the USA, where there is a current policy goal to transform education into an evidence-based field. The US government Strategic Plan for Education asserts that, unlike medicine, education operates:

> 'largely on the basis of ideology and professional consensus. As such, it is subject to fads and is incapable of the cumulative progress that follows from the application of the scientific method and from the systematic collection and use of objective information in policy making' (US Department of Education, 2002: 59).

As this harsh observation suggests, there are specific plans in the US to raise the quality of government-funded research. 'Quality' is to be judged by seeking the opinion of an independent review panel of qualified scientists

and by increasing the percentage of randomized experimental designs when causal questions are addressed. There is a five-year plan in progress.

The quality of educational research in the UK has also come in for criticism from influential sources (Hargreaves, 1997; Tooley and Darby, 1998; HERO, 2002) and there have been recent policy initiatives to strengthen the evidence base for education. An independent National Education Research Forum has been set up with a remit to provide strategic direction for educational research across England and to raise the quality, profile and impact of educational research (NERF, 2003). In a parallel initiative a UK-wide Teaching and Learning Research Programme has been implemented by the Economic and Social Research Council, to support and develop educational research. This programme includes a Research Capacity Building Network with a focus on fostering research knowledge and skills. A programme of systematic reviews of the educational research literature is ongoing at the Institute of Education in London (Oakley, 2002). There are similar initiatives in Scotland, with the Scottish Executive Educational Department and the Scottish Higher Education Funding Council funding the development of high-quality research programmes to stimulate policy and practice (SHEFC, 2002).

UK educational initiatives are not confined to 'hard' designs. Instead there has been considerable consultation about what high quality research might mean within education. Following such a consultation exercise, NERF published a research and development strategy that commented upon methodology (NERF, 2001). Unlike the USA and the NHS, the report showed little commitment to any research design or set of methods as an index of quality, suggesting that it was preferable to develop 'expert use of currently available methodologies and for continuing methodological development' (p 10). 'Fitness of purpose' – the use of the best research methodology for the research questions asked – was seen to be a key concept for judging quality. The Research Capacity Building Network also argues for the use of appropriate methodologies, although given the current lack of quantitative expertise in the UK research field it recognizes a clear need to increase the capacity to run large-scale studies (Gorard, 2001: 31). It aims to support high quality research designs with quantitative outcomes, but which integrate qualitative methods when these answer questions appropriately. The Scottish initiative is similarly concerned with applying good social science research methods to support evidence-informed research and practice, without any assumptions about particular methodologies.

Educational researchers are therefore being urged to move closer to the position of the NHS in seeking evidence for effective practice, but are retaining a commitment to pluralistic approaches rather than adopting one research paradigm or set of methods.

What kind of research can be carried out to evaluate SLT work in schools?

As discussed above, systems approaches and small-scale qualitative analyses remain the most fitting ways of answering the question 'How shall we be judged?' in individual service delivery contexts. But there is a lack of evidence for 'context-neutral' therapy procedures that have been shown to be generally useful and transferable across SLTs, schools and children. These would be welcome and indeed, as SLT work in schools moves increasingly to an indirect model, working through classroom assistants as well as teachers (Law *et al.*, 2002), the need to develop and evaluate such interventions increases. The relative convergence between sectors in accepting the desirability of providing good evidence, with the recognition that different methods are needed to inform complex practice, should mean that the climate in schools will be helpful to SLTs who wish to evaluate practice in a systematic manner.

Therapeutic intervention, like teaching, is multifaceted, and it would be pointless to make it less complex in an attempt to evaluate effectiveness. The Medical Research Council (MRC, 2000) recognizes this and suggests that evaluation of complex interventions, such as SLT work in schools, requires several phases of investigation moving up through the levels of evidence. The phases move from a preclinical, theoretical stage through modelling and exploratory trials to a 'definitive' RCT and then on to evaluate long-term implementation in the field. The 'preclinical' phase is used to explore theory to help to design interventions, set workable hypotheses, predict confounders and suggest study designs. The linked modelling stage is used to identify key components of the intervention and the underlying mechanisms through which they should influence outcomes. Computer and economic modelling may be used and qualitative methods such as surveys, observations and interviews can clarify what matters within interventions. Case studies can be carried out at this stage. The aim of this phase is to sort out what features of an intervention are important and how factors interact so that further studies build on firm ground.

The third phase, 'exploratory trials', is used to 'test drive' the factors that would be relevant in a full RCT. This phase develops a replicable intervention and considers how it may be compared to an alternative. It is in this phase that variations in the intervention such as how much therapy is given, or how it is delivered, can be tried out on a small scale. This phase also allows a check on whether the therapy intervention can actually be delivered in a predetermined way by relevant therapists, and how much individual deviation from the planned intervention is acceptable. Here also, factors that might influence a large-scale trial can be worked out, such as how to get children into the study,

how to monitor therapy and how to measure outcomes (with choice of outcome measures requiring a lot of thought). Such exploratory trials will often be group studies and should lead on to the large-scale RCT phase. The RCT compares the fully defined intervention to an appropriate alternative in a replicable study of appropriate statistical power. This RCT will, no doubt, be big and expensive, but must also test an intervention that is realistic and deliverable within service constraints. Costs as well as benefits of the interventions should be calculated. The last phase, 'long-term implementation' requires confirmatory studies, often in new contexts and therefore often carried out by different services.

There are few interventions for school-aged children that have gone through these phases. A rare exception is 'FastForWord-Language'. The theoretical work of Paula Tallal and colleagues resulted in the development of a computer-based adapted speech programme in a pilot form, which was successful in a small-scale introductory trial (Merzenich *et al.*, 1996). An intervention package on CD-ROM was developed by Scientific Learning Corporation, and a series of case studies and exploratory trials suggesting some effectiveness followed (Tallal, 2000; Gillam *et al.*, 2001). These were followed up by an appropriately powered RCT measuring outcomes with children with persistent receptive SLI (Cohen *et al.*, in press). The clearly defined nature of the intervention was a major factor in allowing these phases to take place.

Gorard (2002) argues that the phases identified by the Medical Research Council (MRC) can also be useful in evaluations of complex educational practices, which is further reassurance that educational and health approaches do not inhabit completely different domains of science. Indeed, Gorard's aim of integrating qualitative methods into high quality research designs with quantitative outcomes is relevant. Instead of running sequential studies, as suggested by MRC, health service research projects increasingly use qualitative methods in the context of RCTs to answer questions about client perceptions and how acceptable they found the intervention.

Recognizing that there are research designs that cope with complex interventions and that different research methodologies fit different purposes can be helpful for SLTs working in schools. It can remind us that we are unlikely to go straight to randomized trials in therapy evaluation and that there is value in other approaches. Case study series could and, I argue, should, be set up in schools and could develop as far as controlled cohort studies. There are some widely-used materials and language programmes that would make good starting points for evaluation, where research could be fitted in to existing service practice. Such studies would not control for all confounding variables, but would at least suggest whether intervention was effective. Quantitative measures of children's language and communication outcomes would be needed, but so also would

qualitative methods to evaluate how children, their schools and parents experienced therapy. This would be a major step forward. Teachers are now being encouraged to engage in research (TTA, 2003; and see Rose, 2002 concerning children with additional learning needs) and collaborative approaches are possible. The SLT profession would be considerably enhanced by a series of cohort studies and this is a realistic objective to achieve.

Although moving from cohort studies to RCTs may be daunting and extra funding can be necessary, it is relatively easy to scale up to a full trial once there is 'lower level' evidence of effectiveness; even the problems of random allocation can usually be overcome. Where they cannot, there may be good ethical reasons for avoiding RCTs and, in any case, research studies of any type must always be submitted for approval to NHS ethics committees in line with the practices of research governance. There is a great deal of advice available about how to construct research studies, especially from medical statisticians who are employed by most NHS trusts. There is also help and support for researchers in education of all kinds on the RCBN website at http://www.cf.ac.uk/socsi/capacity/index.html

Conclusion

SLTs can already provide evidence of their effectiveness in individual school contexts, as measured by children's attainments of classroom and therapy targets and by reports of good collaboration. There remains however a large need for studies of all kinds detailing in a replicable format what interventions are offered to children, and studies which suggest that these interventions work using validated scientific techniques. There is a demand for such studies to meet the NHS objective of offering the best interventions available. Educational researchers are addressing similar issues and the two sectors are developing complementary procedures to evaluate complex interventions.

School-based SLT services should be in a good position to develop and test therapy interventions and do not have to jump straight to RCTs. Case study series and cohort studies appear to be the designs most likely to be achievable and these can be justified as necessary steps on the way to constructing a sound body of evidence. Building up such a body of validated intervention practices would be professionally valuable to SLTs, as well as helping to meet the needs of many children and allowing indirect therapy approaches to be introduced with more confidence of success. It is daunting to be embarking on the 21st century with so few validated therapy practices available for school-

aged children. Despite the complexities involved, it is probably time to demonstrate that therapy in schools works.

Acknowledgements

References

Boyle, J., McCartney, E., O'Hare, A., Forbes, J. and Beirne, P. 1999: *A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment.* Proposal to the UK National Co-ordinating Centre for Health Technology.

Centre for Evidence-Based Medicine 2003: *Study design.* http://www.cebm. net/study_designs.asp (accessed 28 February, 2003).

Clarke, M. and Oxman, A. D. 2003: *Cochrane reviewers' handbook* [updated January 2003]; Section 4.1.6. http://www.cochrane.dk/cochrane/handbook/handbook.htm (accessed 31 March, 2003).

Cohen, W., Hodson, A., O'Hare, A., Boyle, B., Durani, T., McCartney, E., Mattey, M., Naftalin, L. and Watson, J. 2003: Effects of computer-based intervention using acoustically modified speech (FastForWord Language[®]) in receptive language impairment: outcomes from a randomised controlled trial. *Journal of Speech, Language and Hearing Research*, in press.

Gersten, R., Baker, S. and Lloyd, J. W. 2000: Designing high quality research in special education: group experimental design. *Journal of Special Education* **34**, 2–18.

Gillam, R. B., Frome Loeb, D. and Friel-Patti, S. 2001: Looking back: a summary of five exploratory studies of FastForWord. *American Journal of Speech-Language Pathology* **10**, 269–73.

Gorard, S. 2001: A changing climate for educational research? The role of research capacity building. *Occasional Paper Series 45.* Cardiff: Cardiff University School of Social Sciences – ESRC/TLRP.

Gorard, S. 2002: How do we overcome the methodological schism (or can there be a 'compleat' researcher)? *Occasional Paper Series 47.* Presentation to the British Educational Research Association: Exeter, 12–14 September, 2002.

Hammersley, M. 2001: *Some questions about evidence-based practice in education.* Presentation to the British Education Research Association: Leeds, 13–15 September, 2001.

Hargreaves, D. H. 1997: In defence of research for evidence-based teaching: a rejoinder to Martyn Hammersley. *British Educational Research Journal* **23**, 405–19.

HERO 2002: *Research Assessment Exercise 2001: overview of reports from the panels – education.* http://www.hero.ac.uk/rae/overview (accessed 1 April 2003).

HMIE 2002: *How good is our school? Self-evaluation using quality indicators.* www.hmie.gov.uk/documents/publications/HGIOS.pdf. Her Majesty's Inspectorate of Education (accessed 2 April, 2003).

Jones, M., Gebski, V., Onslow, M. and Packman, A. 2001: Design of randomised controlled trials – principles and methods applied to a treatment for early stuttering. *Journal of Fluency Disorders* **26**, 247–67.

Law, J., Boyle, J., Harris, F., Harkness, A. and Nye, C. 1998: Screening for speech and language delay: a systematic review of the literature. *Health Technology Assessment* **8**, 1–184.

Law, J., Lindsay, G., Peacey, N., Gascoigne, M., Soloff, N., Radford, J. and Band, S. 2002: Consultation as a model for providing speech and language therapy in schools: a panacea or one step too far? *Child Language Teaching and Therapy* **18**, 145–63.

McCartney, E. 1999: *Speech/language therapists and teachers working together.* London: Whurr.

McCartney, E. and van der Gaag, A. 1996: How shall we be judged? Speech and language therapists in educational settings. *Child Language, Teaching and Therapy* **12**, 314–27.

McCartney, E., Boyle, J., Bannatyne, S., Jessiman, E., Campbell, C., Kelsey, C., Smith, J. and O'Hare, A. 2004: Becoming a manual occupation? The construction of a therapy manual for use with language impaired children in mainstream primary schools. *International Journal of Language and Communication Disorders* **34**, 135–148.

McCartney, E., Boyle, J., Ellis, S. and Turnbull, M. 2003: *A survey and cohort intervention using indirect speech and language therapy for children with primary language impairment in schools.* Proposal to the West of Scotland Research and Development Project.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L. and Tallal, P. 1996: Temporal processing deficits of language-learning impaired children ameliorated by training. *Science* **271**, 77–80.

MRC 2000: A framework for development and evaluation of RCTs for complex interventions to improve health. London: MRC.

Methany, N. and Panagos, J. 1978: Comparing the effects of articulation and syntax programs on syntax and articulation improvement. *Language, Speech and Hearing Services in Schools* **9**, 50–56.

NERF 2001: *A research and development strategy for education: developing quality and diversity.* Nottingham: National Education Research Forum.

NERF 2003: National Education Research Forum terms of reference. http://www.nerfuk.org/terms.html (accessed 2 April, 2003).

Oakley, A. 2002: Social science and evidence-based everything: the case of education. *Educational Review* **54**, 277–86

Robson, C. 2002: *Real world research.* (2nd edn). Oxford: Blackwell

Rose, R. 2002: Teaching as a 'research-based profession': encouraging practitioner research in special education. *British Journal of Special Education* **29**, 44–48.

Sackett, D. L., Straus, S. E., Scott Richardson, W., Rosenberg, W. and Haynes, R.B. 2000: *Evidence-based medicine.* Edinburgh: Churchill Livingstone.

SHEFC 2002: Applied educational research. Joint scheme by the Scottish Executive Educational Department and the Scottish Higher Education Funding Council. http://www.shefc.ac.uk/content/library/circa/02/he2602/he2602a.htm (accessed 1 April, 2003).

SOEID 1999: *Raising standards – setting targets. Targets for pupils with special educational needs.* February 1999. The Scottish Office http://194.200.85.106/library/documents-w&sten-oo.htm (accessed 14 December, 1999).

Tallal, P. 2000: Experimental studies of language learning impairments: from research to remediation. In Bishop, D. V. M. and Leonard, L. B. editors, *Speech and language impairments in children: causes, characteristics, intervention and outcome.* Philadelphia, PA: Psychology Press Ltd.

Tooley, J. and Darby, D. 1998: *Educational research: a critique.* London: Office for Standards in Education.

Torgerson, C. J. and Torgerson, D. J. 2001: The need for randomised controlled trials in educational research. *British Journal of Educational Studies* **49**, 316–28.

TTA 2003: Promoting research and evidence informed practice (REIP). The Teacher Training Agency http://www.canteach.gov.uk/publications/community/research/evidence/tta00_03.pdf (accessed 18 March, 2003).

US Department of Education, 2002: Strategic plan 2002–2007. http://www.ed.gov/pubs/stratplan2002–07/stratplan 2002–2007.pdf (downloaded 7 March, 2003).